The functional genetics of respiratory syncytial virus induced bronchiolitis : an investigation into the transcriptional regulation of *IL8* and *LTA*

Thesis

TITLE PAGE

Name: Dr Doug Hacking

Degrees: BMSc. Biochemistry (First Class Honours)

Dundee University

BM BCh

Oxford University

Thesis Title: The functional genetics of respiratory syncytial virus induced bronchiolitis: An investigation into the transcriptional regulation of *IL8* and *LTA*

Degree for which thesis submitted: PhD

Sponsoring Establishment: Weatherall Institute of Molecular Medicine,

John Radcliffe Hospital, Oxford

Collaborating Establishment: Wellcome Trust Centre for Human Genetics,

University of Oxford, Oxford

Department of Paediatrics,

University of Oxford,

John Radcliffe Hospital, Oxford

Submission date: 30 July 2004

Award date: 18 November 2004

INDEX

ABBREVIATIONS

A	Adenine
AP-1	Activator protein-1
C	Cytosine
CI	Confidence interval
CTP	Cytosine Triphosphate
C/EBP	CAAT-enhancer binding protein
NFIL6	Nuclear factor interleukin-6 (also known as C/EBP)
CD	Cluster designation
cDNA	complementary DNA synthesized by reverse transcriptase
CEPH	Centre d'Etude du Polymorphisme Humain
CLD	Chronic lung disease
CTD	C-terminal domain
c-Myb	*c-myb* proto-oncogene product
CsCl	Cesium chloride
CTL	Cytotoxic T cell lymphocytes
dATP	deoxy-Adenine Triphosphate
DBD	DNA binding domain
dCTP	deoxy-Cytosine Triphosphate
ddATP	dideoxy-Adenine Triphosphate
ddCTP	dideoxy-Cytosine Triphosphate
ddGTP	dideoxy-Guanine Triphosphate
ddTTP	dideoxy-Thymidine Triphosphate
dGTP	deoxy-Guanine Triphosphate
dTTP	deoxy-Thymidine Triphosphate
DMEM	Dulbecco's modified eagles medium
DNA	Deoxyribonucleic acid
EBV	Epstein-Barr Virus
EDTA	Ethylenediaminetetraacetric Acid
ELISA	Enzyme-linked immunosorbent assay
EMSA	Electrophoretic mobility shift assays
F	RSV Fusion protein
$F_{12}K$	Kaighn's modified media
FISH	Fluorescent *in situ* hybridization
G	Guanine
G	RSV Glycoprotein: viral attachment to the cell
GAG	Glycosaminoglycans
GM-CSF	Granulocyte-macrophage colony stimulating factor
GRO	Growth-stimulatory activity/growth-regulated gene
HBD	Heparin binding domain
hnRNP	heterogeneous nuclear ribonucleoprotein
HR	Heptad repeat sequences
ICAM	Intracellular adhesion molecule
Ig	Immunoglobulin
IQR	inter-quartile range
IL	Interleukin
IFN	Interferon

L	RSV Nucleocapsid protein: RNA polymerase
M	RSV Matrix protein: viral assembly
M2	RSV M2 transcription elongation factor
MALDI-TOF	Matrix-assisted laser desorption time-of-flight mass spectrometry
MAPREC	Mutant analysis by PCR and restriction enzyme cleavage
MCP	Macrophage chemoattractant protein
MHC	Major histocompatibility complex
MIP	Macrophage Inflammatory Protein
mRNA	messenger RNA
MOI	Multiplicity of infection
MPO	Myeloperoxidase
N	RSV Nucleocapsid protein: Nucleoprotein
NO	Nitric Oxide
NF-κB	Nuclear factor-κB
NK	Natural Killer
L-NMMA	N^G-monomethyl-L-arginine
NOS	Nitric Oxide Synthase
NS 1 or 2	RSV Non-structural proteins
NTP	Nucleotide triphosphate
OXREC	Oxfordshire research ethics committee
P	RSV Nucleocapsid protein: Phosphoprotein
PCR	Polymerase chain reaction
PE/MS	Primer extension and Mass spectrometry
PEG	Polyethylene glycol
PF-4	Platelet factor-4
PMA	4α-Phorbol 12-myristate 13-acetate
PML	Polymorphonuclear neutrophils
Pol	Polymerase
PPBP	Pro-platelet basic protein
pRSV	Purified respiratory syncytial virus
RANTES	Regulation upon activation normal T cell-expressed and secreted
REL	Reticuloendotheliosis viral oncogene homolog/NF-κB family
RSV	Respiratory syncytial virus
RSVRE	RSV response element
RNA	Ribonucleic acid
RNI	Reactive Nitrogen Intermediates
SH	RSV Small hydrophobic protein
SNP	Single nucleotide polymorphism
SP	Surfactant protein
T	Thymidine
TDT	Transmission/disequilibrium test
TE	Tris-EDTA
TH	T helper type lymphocyte
TLR	Toll-like receptor
TNF	Tumor necrosis factor-α
upRSV	Unpurified respiratory syncytial virus
UV	Ultraviolet

ABSTRACT

This thesis explores the extent to which natural variation in DNA sequence may influence the transcriptional regulation of genes that play a critical role in the immune response to infection. As a model system we began by examining Interleukin-8 (*IL-8*) as it has been implicated in the pathogenesis of a very common infection namely, RSV-induced bronchiolitis. Previously, we have described an association between bronchiolitis disease severity and a specific *IL8* haplotype comprising six single nucleotide polymorphisms (SNPs) (-251A/+396G/+781T/+1238delA/+1633T/+2767T, haplotype 2). Here we investigated the functional basis for this association by measuring haplotype-specific transcription *in vivo* in human primary cells. We found a significant increase in transcript level derived from the *IL8* haplotype 2 relative to the mirror haplotype 1 (-251T/+396T/+781C/+1238insA/+1633C/+2767A) in respiratory epithelial cells but not in lymphocytes. A promoter polymorphism, -251A, present on the high producer haplotype, had no clear and significant effect on the allele-specific level of transcription when analyzed in reporter gene experiments in human respiratory epithelial cell lines. We proceeded to systematically screen for allele-specific protein-DNA binding in this functional haplotype which revealed significant differential binding at the +781T/C polymorphism. C/EBP β was identified as being part of a transcription factor binding complex which preferentially bound in the presence of +781T allele. These results suggest that the mechanism for disease susceptibility to RSV-induced bronchiolitis may occur through a haplotype-specific increase in *IL8* transcription which may be mediated by functional polymorphisms within that haplotype.

However, the unusual mirror haplotype structure of *IL8* in European populations gave little indication of which SNP within the haplotype was likely to be functional *in vivo*. In order to demonstrate the principle that a non-coding polymorphism may alter transcriptional regulation *in vivo* I used a second technique, namely the haplotype-specific chromatin immunoprecipitation (haploChIP), which once optimised could be applied to the *IL8* locus. In order to optimise haploChIP I studied a second model system namely the tumor necrosis factor/lymphotoxin-α (*TNF/LTA*) locus and used it to measures the differences in the loading of RNA polymerase between alleles in EBV-immortalized B lymphocytes. Consistent with previous studies I demonstrated that allele-specific RNA polymerase loading was independently associated with the *LTA*+80, *LTA*+368 and *LTA*+723 polymorphisms. This association was refined to demonstrate that the polymerase loading at this locus appeared to be determined *in vivo* by the *LTA*+80A and *LTA*+368C alleles both of which occur on the same haplotype. These data imply that HaploChIP could be utilised in understanding the transcriptional regulation of *IL8*. Furthermore it shows that functional SNPs may act in concert to influence transcription. This may mean that in future the functional basis for allele-specific effects at critical candidate genes will have to be dissected *in vivo* at the level of the haplotype.

CHAPTER ONE

INTRODUCTION

1.1

Respiratory Syncytial Virus (RSV) infection in childhood

The majority of deaths from RSV-bronchiolitis in children occur under 1 year of age and in infants with specific risk factors for severe bronchiolitis such as congenital heart disease, chronic lung disease (CLD) or prematurity (Shay et al. 2001). However, the known risk factors associated with RSV-induced bronchiolitis mortality do not provide a full explanation for the variation in disease severity across the whole population. By the age of two years most infants will have encountered the virus and the majority will have had a mild illness with only 2% requiring hospital admission (Medical Research Council 1978). The reasons why some infants are more severely affected by RSV infection than others remains incompletely understood.

In this introduction I will discuss the rationale for a molecular genetic approach using a candidate gene study design. I will go on to examine the biology of the virus and the host response to infection. I will proceed to focus on the role that chemokines, and interleukin-8 (IL-8) in particular, have in the pathogenesis of RSV infection. Then, I will describe the evidence for the association between RSV-induced bronchiolitis disease severity and natural DNA sequence variation at or close to the *IL8* gene.

1.2.1

Rational for the molecular genetic approach

Dissecting out the critical events in RSV bronchiolitis that lead to enhanced disease severity is hampered by four principal problems. Firstly, there are a large number of mediators produced by the host in the event of an infection each of which has the potential to be protective or deleterious. Secondly, animal experiments may prove to be misleading, as there may be critical differences in the pathogenesis between species. For instance, the experiments on RSV bronchiolitis in mice suggest that IL-4 activation of mast cells and IL-5 mediated chemotaxis of eosinophils may be crucial to the pathogensis of the disease (Graham 1995). However, in humans the predominant leukocyte involved in the immune response to RSV is the neutrophil (Everard et al. 1994). Thirdly, clinical research may help to generate hypotheses through descriptive studies but may not be able to test these hypotheses or demonstrate causality of a particular mediator in the pathogenesis of the disease. Finally, inter-species genomic sequence comparisons have proved to be a powerful means of identifying mammalian *cis*-acting regulatory elements (Pennacchio and Rubin, 2001). At the time this thesis was conducted publicly available genome sequence data was only available for the mouse. Unfortunately, in the case of *IL8,* human to mouse inter-species comparisons would be of limited value as this chemokine is absent in mice (Modi and Yoshimura, 1999).

There is no clear evidence to suggest a particular RSV subgroup is associated with disease severity (Fletcher et al. 1997; Smyth et al. 2002). However, a number of studies of inflammation have demonstrated a difference in cytokine production, namely tumor necrosis factor-α (TNF), IL-10, and IL-1β, between individuals (Westendrop et al. 1997; Turner et al;

1997, Bioque et al. 1995) which may be related to the increasing number of DNA polymorphisms being described located in the coding or promoter regions of these genes. The polymorphisms in the gene may modify the protein product of the gene. For instance, there is evidence that a polymorphism (Δ32 deletion) in the chemokine CCR5 gives selective advantage in resisting human immunodeficiency virus infection (Murphy 2001). Alternatively, polymorphisms in the promoter region may alter the binding affinity of transcription factors (Knight et al. 1999) which in turn could change the level of expression of a gene. Therefore, investigation of the functional properties of polymorphisms associated with increased susceptibility to RSV-induced bronchiolitis may serve to explain both the variable clinical severity of the disease as well as demonstrate the causal mechanisms of its pathogenesis.

1.2.2

Association studies with candidate genes

Association studies using candidate genes have greater power to detect genetic factors involved in complex disease than other methods such as linkage analysis (Risch and Merikngas, 1996). Candidate gene studies focus on genes for which there is an *a priori* hypothesis about their role in the aetiology of a disease and are conducted in a population sample of affected and unaffected individuals. Therefore, a candidate gene study can be seen to take advantage of both the increased statistical efficiency of association analysis in a complex disease as well as the biological understanding of the genes and proteins likely to be involved its pathogenesis (Tabor et al 2002). However, studies of candidate genes have been criticised as the significance of an association is often either overstated in the original report when compared to subsequent studies or is found not to have been replicated at all

(Gambaro et al. 2000; Ioannidis et al. 2001). Carefully conducted candidate gene studies can yield reliable results especially when the biological plausibility of the association is strong, the strength of the association is great and when the association is consistent across different populations. Moreover, a true association is more likely when a dose response relationship of the association is seen such that an individual with two copies of the disease associated variant is at greater risk than an individual with only one. The important issue of unrecognized population stratification, leading to a spurious association, can also be avoided by careful study design and through the use of family based studies where parents act as biological controls for their affected children (Reider et al. 1999).

Viral Biology

1.3.1

An introduction to RSV

Human respiratory syncytial virus (RSV) was first characterized in 1957 (Chanock et al. 1957) and has since been recognized as the most common cause of severe respiratory tract infection in young infants worldwide. RSV infection is not restricted to children as immunity following infection remains incomplete and secondary infections can occur throughout life. Despite 45 years of research there is neither an effective treatment nor any immediate prospect of a vaccine. Key to the development of novel therapies is a good understanding of RSV disease pathogenesis.

1.3.2

RSV structure and genome

RSV is classified in the genus *Pneumovirus* of the family *Paramyxoviridae* (Figure 1.1). The RSV genome is composed of single stranded negative sense RNA of 15,200 nucleotides encoding 11 proteins. The structure of the RSV genome and protein functions are given in detail in Figure 1.2.

The ability to reconstitute infective RSV particles by introducing engineered viral genes into cells was a break-though in RSV research and is responsible for much of our understanding of the function of the proteins in the virus. In this technique the RSV genome is isolated and converted from RNA to DNA. This DNA can then be cloned into vectors called plasmids which allows a large number of copies of the genome to be made. Plasmids

encoding the viral proteins as well as the entire viral genome are then introduced in to epithelial cells where they are transported to the nucleus. The cell's own machinery then makes all the necessary proteins for viral assembly and finally replicates the RSV genome which is then packaged into the viral particle. Mutations of the DNA in the plasmids encoding various RSV genes can then be introduced, including deleting entire genes, and their effects on viral replication both *in vitro* and *in vivo* can be observed. For example, use of this system has shown that the genes NS1, NS2, SH, M2-2 and G can be mutated whilst maintaining viral viability (Schmidt et al. 2001). However, deletion of the NS1 or NS2 genes results in a marked reduction in replication efficacy *in vivo* (Whitehead et al. 1999, Schlender et al. 2000) a difference which has been attributed to the role these genes have in co-operatively antagonizing α and β interferon (Teng et al. 2000).

As well as leading to discovery of the function of RSV proteins, this system has been used to develop vaccine candidates. An ideal RSV vaccine would be highly attenuated (that is, causes no illness), but retain high replication efficiency (so that it can be given as a single intra-nasal dose) and high immunogenicity. Unfortunately to date these vaccines have proved insufficiently attenuated to be introduced to the population.

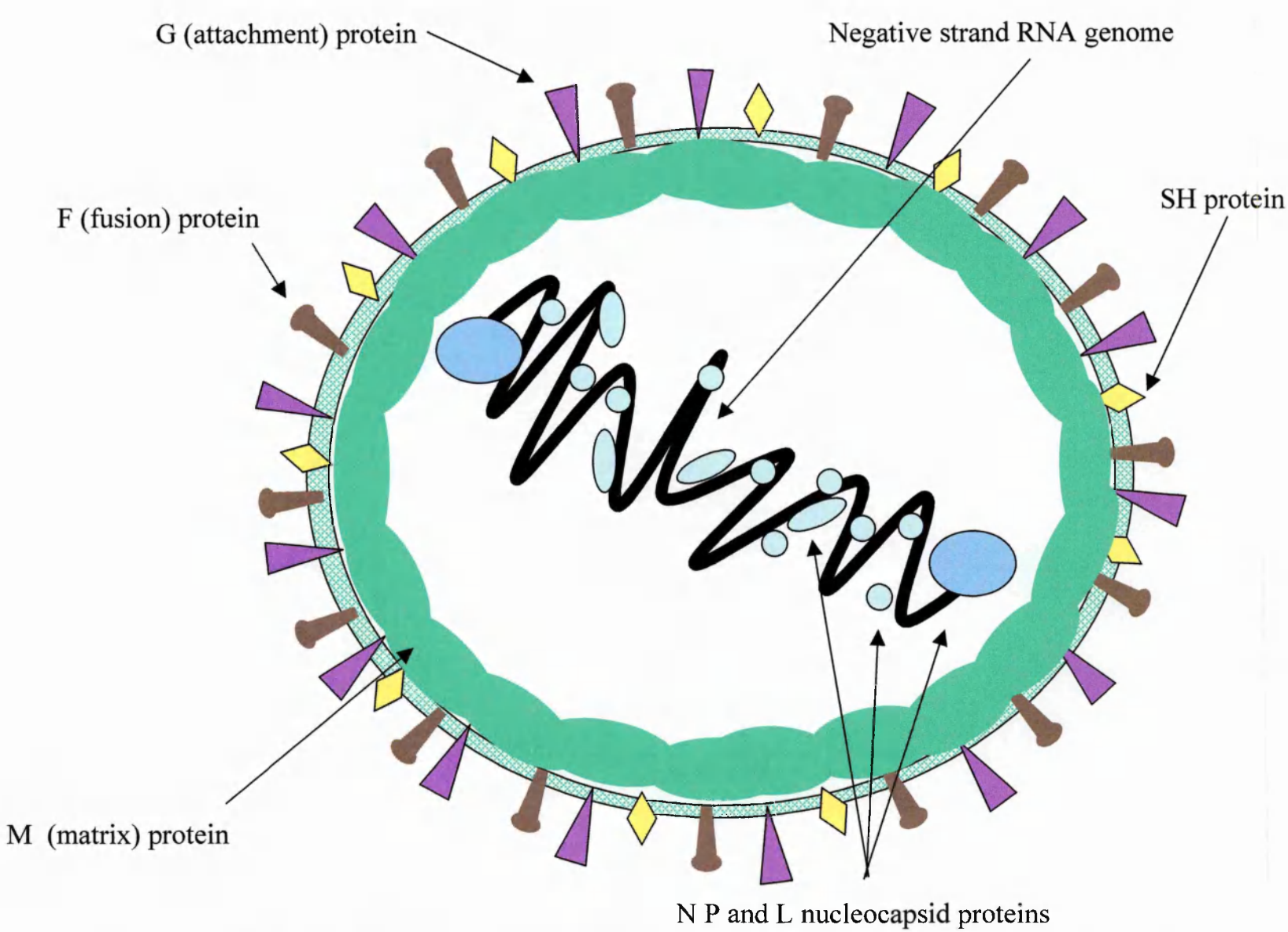
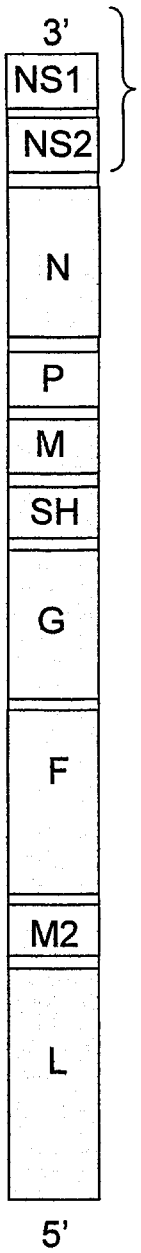


Figure 1.1. Respiratory syncytial viral structure. The RSV virion is about 200 nm in size and consists of a nucleocapsid within a lipid envelope. The nucleocapsid is a symmetrical helix with a helical diameter of 12 to 15 nm. The lipid bilayer is derived from the host plasma membrane and contains virally encoded transmembrane surface glycoproteins which are 11 to 20nm in size and closely spaced at intervals of 6 to 10 nm.



Protein	Function
NS1 NS2	Non-structural proteins: anti-interferon α and β activity
N	Nucleocapsid protein: Nucleoprotein essential for transcriptional activity
P	Nucleocapsid protein: Phosphoprotein essential for transcriptional activity
M	Matrix protein: viral assembly
SH	Small hydrophobic protein: function unknown
G	Glycoprotein: viral attachment to the cell
F	Fusion protein: viral entry and syncytia formation
M2	M2-1: transcription elongation factor M2-2: regulation of viral transcription
L	Nucleocapsid protein: RNA polymerase

Figure 1.2. The RSV genome, proteins and protein function. The RSV genome is a single stranded negative sense RNA of 15,200 nucleotides that is transcribed into 11 subgenomic mRNAs. Transcription is initiated from the 3' end with only a fraction of the polymerase moving on to the next gene. This creates a gradient of transcriptional attenuation with distance from the transcriptional start site reflecting the required relative abundance of the encoded proteins (Collins et al. 1996).

1.3.3

RSV Infection of Airway Epithelial Cells

RSV predominantly infects airway epithelial cells. These cells line the nose as well as the large and small airways. They form the first line of defense against the virus and are the site of the majority of the inflammation associated with the disease. In immuno-competent individuals the virus seldom infects other tissues. The first critical step in the infection process is entry of the virus into the cell. Thereafter the virus replicates, is packaged and leaves the cell, either through fusion with adjacent cells or following cell rupture (Figure 1.3). Blocking viral entry is an attractive mechanism for anti-viral therapies and for this reason most research work has been directed towards the functions of the major surface glycoproteins F and G. Their importance is emphasized by the observation that effective humoral immunity is only conferred by antibodies to these proteins.

1.3.4

Role of the G protein

The G glycoprotein was originally identified as the RSV attachment protein in experiments in which binding of radio-labeled RSV particles to cultured cells was blocked with specific anti-G antibodies (Levine et al. 1987). The G protein is a glycosylated type II transmembrane protein of between 289 to 299 amino acids depending on the viral strain. It consists of an N terminal cytoplasmic domain, a hydrophobic anchor region and an ectodomain. The ectodomain is made up of two mucin-like regions which are highly variable and rich in serine, threonine and proline. These mucin-like regions are also heavily glycosylated with *N*- and *O*- linked sugars. The mucin-like domains are separated by a 13 amino acid sequence which contains four cystine residues that are conserved across RSV

strains. Many of the epitopes recognized by the host antibody response lie in the C terminal variable region. The virus may have the potential for changing the antigenic profile of this region of the G protein in order to evade the host response (Cane 2001). *N*-glycosylation that masks the epitopes could be another means of evading immune recognition of antigens (Cane 1997). In addition to the membrane bound G protein an N-terminal truncated soluble and secreted form is produced. It was suggested that the soluble G protein served to redirect the immune response away from infected cells and released virus (Cane 2001). However, recent experiments have shown that RSV strains lacking the membrane bound G protein but expressing the soluble G protein had efficient replication *in vitro* and were only slightly attenuated *in vivo* (Teng et al. 2001). This may mean that the soluble form can remain associated with the virus in order to mediate cell entry. Surprisingly, recombinant RSV in which the secreted and membrane bound G protein have both been deleted are still able to infect some cell types as efficiently as wild-type virus. This mutant form of the virus has a much lower efficiency of infection of human airway cells and in the airways of mice *in vivo* (Teng et al 2001, Karron et al. 1997). These results indicate that the G protein is not essential for cell attachment, but acts as an accessory protein which increases the efficiency of the process (Teng et al. 1998).

The cellular receptor for the G protein has not been identified. The cell surface glycosaminoglycans (GAGs) heparan sulfate and chrondroitin sulfate B have been shown to be important for infection *in vitro* through a putative heparin binding domain (HBD) (184-198 amino acids) in the G protein (Feldman et al. 1999). However, RSV with a mutation of the HBD remained sensitive to neutralization with soluble heparin *in vitro*, had normal growth *in vitro* and showed only a mild reduction *in vivo* growth (Teng et al. 2001). This

implies that the HBD is not the sole determinant of heparin sensitivity. Interestingly, the same HBD region of the G protein may prove to have other functions as it contains a motif at 182-186 amino acids common to the CXC3 chemokine fractalkine (Johnson et al. 1987). Tripp and colleagues have gone on to show that the G protein binds the CX3C receptor (CX3CR1) and in doing so can facilitate RSV infection (Tripp et al. 2001). It is possible that by competing with fractalkine for the CX3C receptor, RSV is able to alter fraktalkine mediated immune responses and favor viral replication and survival.

1.3.5

Role of the F protein

RSV has an absolute requirement for the F protein (Teng et al. 1998). Fusion of the viral envelope or infected cell membranes with uninfected cell membranes is an essential step in the virus life cycle. This process requires the F protein to be intact and is enhanced by the G protein. The F protein is a highly conserved protein among the *Paramyxoviridae* family (Collins et al. 1996) and is synthesized as a 67 kDa precursor which undergoes proteolytic cleavage to produce two disulfide-linked polypeptides named F1 from the C terminus and F2 from the N terminus (Collins and Mottet 1991). The part of the F protein that enters the cell membrane is situated at the N terminus of the F1 polypeptide, whilst the transmembrane segment is located close to the C terminus. Adjacent to these two regions are two heptad repeat sequences, denoted HR-N and HR-C, that form stable trimer hairpin-like structures that undergo a conformational change to enable the viral and cell membranes to be apposed prior to viral entry. The heptad structures may prove to be a successful target for drug development as viral fusion can be inhibited by peptides that bind to the HR-C region and

prevent the conformational change necessary for membrane fusion (Lambert et al. 1996). The F protein also interacts with a Ras family protein called RhoA (Patsey et al. 1999). RhoA is a small GTPase that has numerous roles in the cell which include IL-1β, IL-6 and IL-8 secretion, cell motility, cell morphology and most importantly actin cytoskeleton reorganization. Whilst RSV F and RhoA interaction in the infected cell membrane has yet to be demonstrated, RhoA is activated on RSV infection (Gower et al 2001) and RhoA expression facilitates syncytia formation in RSV infected HEp-2 cells over-expressing RhoA (Patsey et al. 1999). Moreover, a RhoA derived peptide bound to the F protein at high affinity prevented cell free viral entry and syncytia formation *in vitro* and infection in mice (Patsey et al. 2000). Another means of inhibiting RSV induced RhoA activation is by preventing RhoA association with the cell membrane following GTP exchange and isoprenylation. Accordingly isoprenylation inhibitors, such as lovastatin, have been shown to have anti-RSV activity both *in vitro* and *in vivo* (Gower and Graham 2001).

1.3.6

RSV assembly

The M protein is vital for RSV viability and plays a central role in RSV assembly (Figure 1.3) and through a series of interactions it co-ordinates the assembly of the envelope proteins F and G with the nucleocapsid proteins N P and M2-1 (Schmidt et al. 2001).

Figure 1.3. A schematic representation of the replicative cycle of RSV.

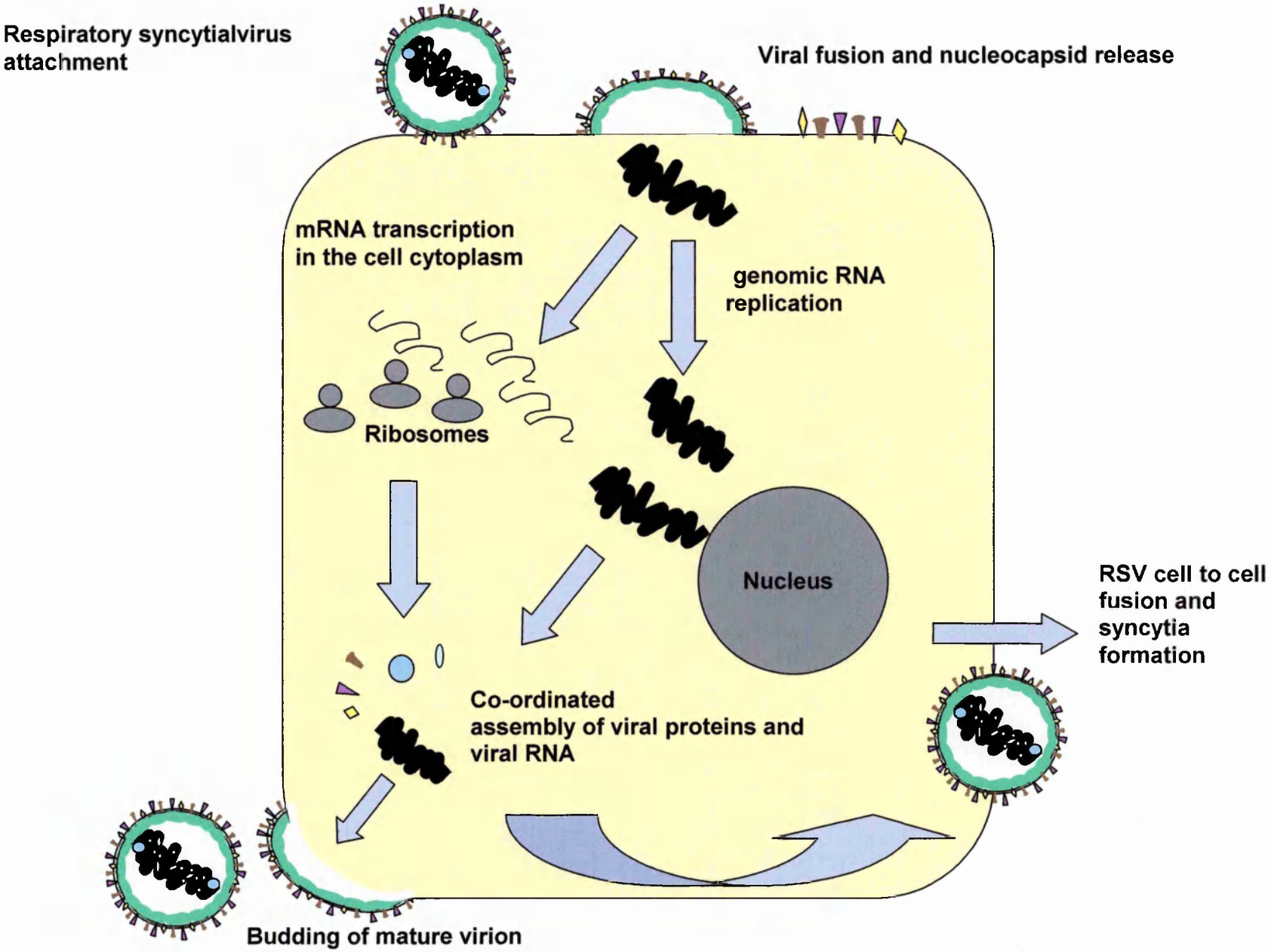


Figure 1.3.

RSV attachment occurs via the G protein. Fusion is mediated by the F protein after which the viral envelope is incorporated into the cell membrane and the nucleocapsid is released into the cytoplasm. RSV mRNA accumulates until about 15 hours after infection and then remains at constant levels thereby allowing for RNA replication and subsequent viral assembly. The M2-2 gene governs this transition from transcription to production of genomic RNA. The M protein and membrane-destined G protein meet in the Golgi and can interact through a six amino acid motif in the cytoplasmic end of the G protein. It is likely that the F and G proteins interact with each other via their cytoplasmic domains. Meanwhile the N, P, L and M2-1 proteins form inclusions in the cytoplasm and then through M2-1 interact with the M protein. Through these series of interactions the M protein can be seen to co-ordinate the assembly of the envelope proteins F and G with the nucleocapsid proteins N P and M2-1. Budding appears to be the reverse of penetration and occurs *in vitro* on the apical cell surface.

Host Response

1.4.1

The host response to RSV infection

The immune response to RSV infection, as with any other infection is comprised of an innate response, and subsequent activation of humoral and cellular specific immunity. Autopsy studies of infants who died following bronchiolitis show lymphocytic infiltrate around small airways and cell debris in airway lumens (Aherne et al. 1970). Broncho-alveolar lavage samples of ventilated RSV infected infants show marked neutrophil infiltrates in the airways (Everard et al. 1994). Moreover, these neutrophils adhere to RSV infected epithelial cells and become activated thereby causing cytotoxic damage to the epithelium (Wang et al. 1998). The cellular immune response seems to be important in controlling the infection once initiated and for clearance of the virus. Children with cellular immunodeficiencies have prolonged viral carriage (Chandwani et al. 1990). The importance of the host response to RSV infection was dramatically illustrated by the formalin-killed RSV vaccine trial in the late 1960s. Vaccinees in this trial made an ineffective response to the vaccine and developed more severe disease during subsequent natural infection (Kim et al. 1969). The humoral immune response is important in providing protection against subsequent infection.

1.4.2

Innate Immunity to RSV

Respiratory epithelial cells whilst being the principle target for RSV infection are also the first defense in the innate immune response to virus (Figure 1.4). Not only do they produce opsonins and collectins but also secrete a range of cytokines such intracellular adhesion molecule (ICAM)-1 (Patel et al. 1995) IL-8 (Noah et al. 1995) and RANTES (Regulation

upon activation normal T cell-expressed and secreted) (Becker et al. 1997), that have been demonstrated to initiate neutrophil, CD4+ T helper cell and eosinophil chemotaxsis. In addition to the action of epithelial cells alveolar macrophages play an important role in the innate immune defense against RSV as they may regulate the immune response (Midulla et al. 1993), and release inflammatory cytokines when infected (Becker et al. 1991). It has recently been shown that the F protein of RSV binds to the Toll-like receptor (TLR) 4 and the CD14 receptor on human monocytes stimulating the release of cytokines IL-1β, IL-6, IL-8 and TNF (Kurt-Jones et al. 2000). Moreover, mice deficient in TLR4 had reduced CD14(+) cell pulmonary trafficking, deficient NK cell function, and impaired RSV clearance (Haynes et al. 2001).

Amongst the collectins produced by epithelial cells surfactant A (SP-A) plays an important role in clearance of RSV. *In vitro* experiments have shown that SP-A binds and neutralizes RSV (Ghildyal et al. 1999) whilst mice deficient in SP-A have impaired clearance of the virus (LeVine et al. 1999). A possible explanation for this delayed viral clearance could be impaired uptake of the virus by mononuclear cells and macrophages as SP-A enhances opsonisation of RSV by these cell types (Barr et al. 2000) thereby increasing TNF and IL-10 production.

Figure 1.4. A schematic representation of RSV infection of the respiratory epithelium

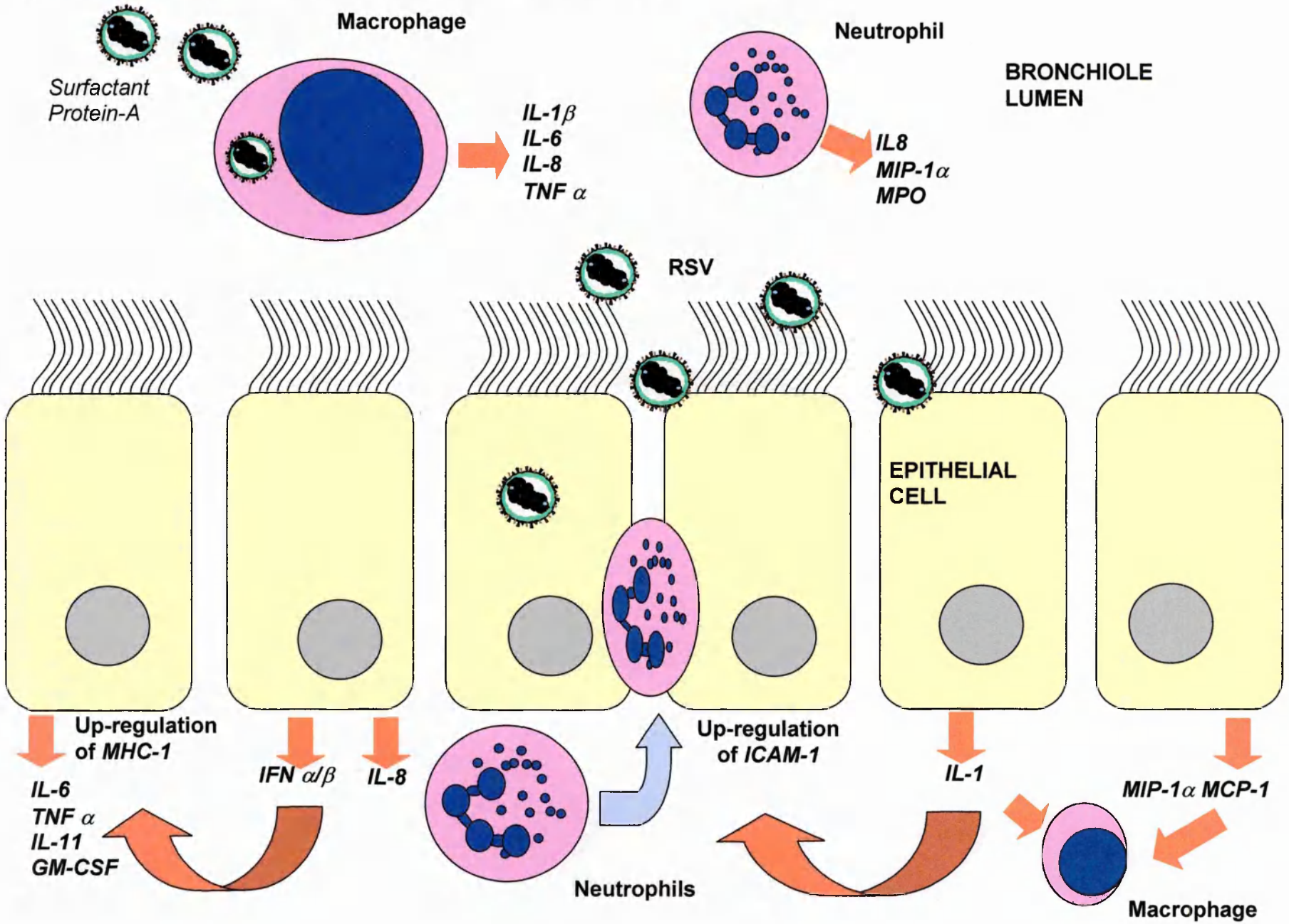


Figure 1.4.

RSV is bound by surfactant protein A and engulfed by macrophages which leads to IL-1β, IL-8, IL-6 and TNF release. Infection of epithelial cells stimulates release of a number of cytokines including the neutrophil and macrophage chemoattractants IL-8 and macrophage chemoattractant protein (MCP)-1. IL-1 secretion leads to the up-regulation of ICAM-1, a receptor vital for neutrophil adherence and activation. IFN α/β secretion causes the up-regulation of the major histocompatibility complex I (MHC-1).

1.4.3

Cell Mediated Immunity to RSV

Infants with a primary RSV infection develop a cellular immune response within 10 days of infection (Chiba et al. 1989). Human cytotoxic T cell lymphocytes (CTL) recognize the N, SH, F, M, M2, and NS2 proteins but not the G protein (Cherrie et al 1992). In mouse models it has been suggested that the failure of the G protein to induce a CTL response is a significant factor in disease pathogenesis (Srikiatkhachorn et al. 1997). Both sensitized CD4+ and CD8+ T cells have been demonstrated to have antiviral activity on passive transfer to RSV infected mice. Both cell types reduce pulmonary shedding of the virus but both also caused increased pulmonary damage (Alwan et al. 1992). Further evidence for the role of CTL in RSV pathogenesis came from animal experiments investigating the effects of formalin-inactivated vaccine. Mice immunized with the vaccine produced CD4+ T cells, with a cytokine expression pattern suggestive of a T helper type two (Th2) phenotype whereas natural infection induces T helper type one (Th1) like cytokine expression (Graham et al. 1993). These data suggested that the formalin-inactivated vaccine induced an imbalance in favor of the Th2 allergy associated response over the Th1 delayed hypersensitivity related response. This proved to be an attractive model to explain the post RSV-induced bronchiolitis recurrent wheeze as well as the allergy associated features of bronchiolitis such as the activation of eosinophils (Garofalo et al. 1992) and the presence of RSV specific IgE (Welliver et al. 1981). However, the Th2 and Th1 CD4+ cell subgroups were first described in mice (Mosmann et al. 1989) and the distinction between the two groups may be less clear in humans (Krug et al. 1996). Moreover, there are contradictory reports on Th2 and Th1 cytokine profiles in human RSV-induced bronchiolitis (Roman et al.

1997, van Schaik et al. 1999) with recent studies relating disease severity more to chemokine production than to Th2 cytokine levels (Garofalo et al. 2001).

1.4.3

Humoral Immunity to RSV

Although the humoral response is not likely to be helpful in influencing the course of a primary infection once it has occurred, there is good evidence that protection from subsequent infections is mediated though antibodies. Breast fed infants with high titers of transplacentally acquired antibodies are less likely to succumb to severe bronchiolitis (Holberg et al. 1991) and passive immunization with RSV-specific IgG has been shown to be effective at protecting against hospital admission with RSV infection (Impact-RSV study Group 1998). Not all humoral responses are favorable and RSV-specific IgE has been associated with more severe disease and with subsequent wheezing (Welliver et al. 1980, Welliver et al. 1986).

Chemokine Biology

1.5.1

Chemokines

The position and movement of cells is essential for normal development and immune function. A fundamental mechanism for cell migration which is shared by both primitive and higher complex organisms is the presence of chemoattractants that signal via seven-membrane G protein-coupled receptors. In humans the most important mediators of

leukocyte migration are a subgroup of these G protein-coupled receptors called chemoattractant cytokines otherwise known as chemokines (MacKay 2001).

1.5.2

Chemokine classification

Though the first chemokine to be discovered was platelet factor 4, chemokines were not recognized to have a chemotactic function until IL-8 was characterized (Yoshimura et al. 1987). Similarly, the first chemokine receptor to be cloned was the IL-8 receptor (Holmes et al. 1991). There are now over 40 known chemokines and 20 chemokine receptors which are classified in four families: CC, CXC, C and CX_3C according to the structure of cysteine residues at the NH_2-terminus (Rossi et al. 2001). For instance, IL-8 is classified as a CXC ligand 8 or CXCL8 whilst the IL-8 receptors are named CXCR1 and CXCR2.

1.5.3

Chemokine and IL-8 function

Consistent with the importance of the chemokines to cell movement is the conservation of these molecules across species. The most likely primordial chemokine is stromal cell-derived factor 1 (CXCL12) whilst the putative primordial chemokine receptor is CXC chemokine receptor 4 (CXCR4)-like receptor. Both CXCL12 and CXCR4 are common to humans and *Drosophila* and have roles in development as well as in the immune system (Gerard and Rollins, 2001). For example, where either molecule is mutated in mice there are major defects in cerebellar development, cardiac septum formation and gastrointestinal development.

Within the immune system chemokines have four main functions. Firstly, chemokines initiate leukocyte mobility through the generation of a chemokine gradient, which the migrating cell follows. The chemokine system allows for specificity in cell migration as the expression of a chemokine from an organ will only promote migration of leukocytes expressing the relevant receptor. Secondly, chemokines have a central role in leukocyte-endothelial cell interactions. This is described in the multi-step model which starts with leukocyte adhesion to the vessel wall using selectin molecules and is followed by integrin binding. On binding the integrins undergo rapid conformational changes which initiate firm adhesion and subsequent trans-endothelial migration. Chemokines, and IL-8 in particular, have been demonstrated to have a role in integrin activation (Gerszten et al. 1999). Moreover, the use of chemokine, integrin and selectin molecules in combination for successful leukocyte extravasation allows for further selection of migrating leukocytes by the tissues. Thirdly, chemokines stimulate leukocytes to release inflammatory mediators. For instance, IL-8 stimulates neutrophil granule exocytosis, and also initiates the respiratory burst which allows production of reactive oxygen intermediates. Finally, chemokines stimulate angiogenesis and angiostasis. The CXC chemokines, such as IL-8 and CCL2, have angiogenic properties (Koch et al. 1992) whereas the CXCR3 ligands, for example CXCL10, have angiostatic properties. The biological relevance of angiogenesis and angiostasis could relate to inflammatory responses and tumour rejection respectively. These four properties are used in combination depending on the biological response required. For instance, tumour rejection requires leukocyte recruitment as well as angiostasis whilst allergic inflammation requires leukocyte recruitment as well as inflammatory mediator release.

RSV-induced bronchiolitis and *IL-8*

1.6.1

The associated of genetic polymorphisms at or close to the IL8 gene with RSV-induced bronchiolitis disease severity

The Kwiatkowski laboratory has shown that a common single nucleotide polymorphism (SNP) -251 (A/T) nucleotides upstream of the *IL8* transcription start site is associated with susceptibility to RSV-induced bronchiolitis (Hull et al. 2000). A further nine SNPs in a 7.6 kilobase segment spanning the *IL8* gene and its promoter were identified of which six (-1722insT, -251A, +396G, +781T, +1633T and +2767T) were used to define the haplotype structure (Table 1.1) of the *IL8* locus. Using these SNPs it was shown that in European populations the haplotype structure is unusual as it is dominated by two common haplotypes (Figure 1.5) (Hull et al. 2001). The definition of these haplotypes allowed the IL8 -251A bronchiolitis association to be refined as Haplotype 12, which contains the -251A and +781T alleles and is associated with RSV-bronchiolitis disease susceptibility (transmission/disequilibrium test [TDT] p=0.0008) (Hull et al. 2001).

Table 1.1

Frequencies of the 12 different Haplotypes observed in a European Population

Haplotype Identifier	**Haplotype***	**Frequency in Europeans****
1	delTTTCCA	.52
2	delTTGCCA	...
3	delTTGTTT	.005
4	delTATCCA	.01
5	insTATCCA	.005
6	delTAGCCA	.03
7	insTAGCCA	...
8	delTAGCTT	.005
9	delTAGTTA	.01
10	delTATTTA	.005
11	delTATTTT	...
12	delTAGTTT	.41

*The SNPs comprising these haplotypes (and haplotype frequencies) are: -1722insT (0.01), -251A (0.47), 396G (0.46), 781T (0.43), 1633T (0.44) and 2767T (0.42). **Frequencies are estimated from a sample of 190 European chromosomes.... = not found.

Adapted from Tables 3 and 4 in Hull et al. (2001) and used with permission from the author.

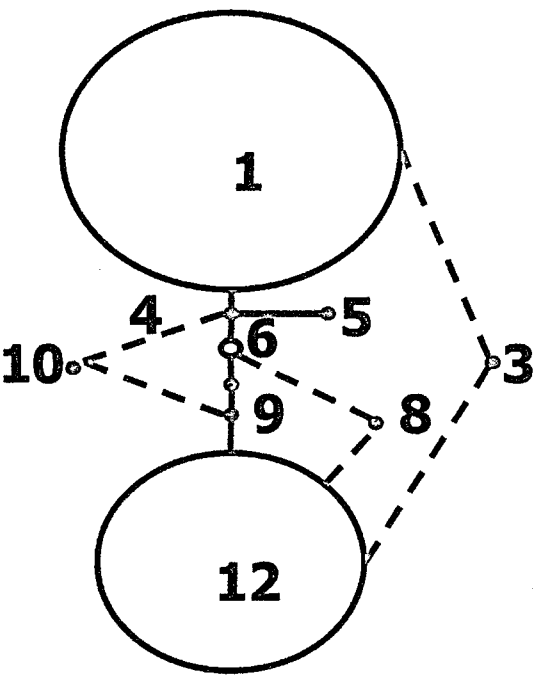


Figure 1.5.

Suggested minimum-mutation network of the *IL8* haplotypes in European populations. Each haplotype is represented by a circle, the radius of which is proportional to its allele frequency. The solid branches (———) represent mutation events and the dashed branches (- - - -) represent recombination events. In each genealogy, the haplotype refers to haplotypes shown in table 1.1. Sequence data from the chimpanzee suggests that the ancestral root is haplotype 4. Adapted from Figure 1 in Hull et al. (2001) and used with permission from the author.

1.6.2

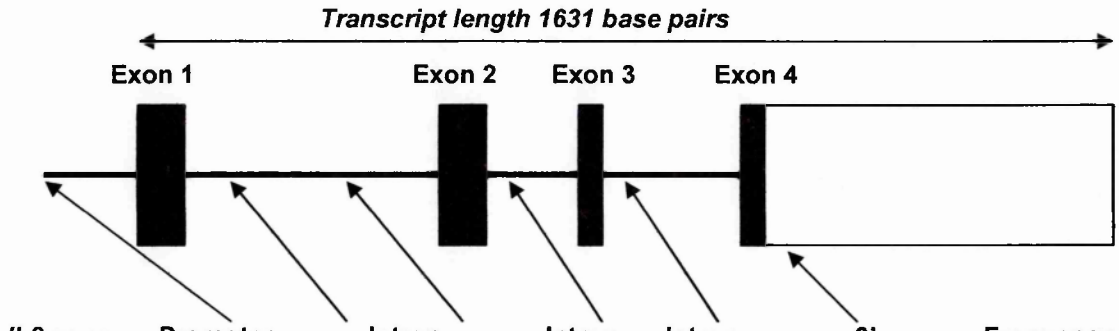
Determining functional genetic elements that may explain the association between the IL8 gene and RSV-induced bronchiolitis disease severity

The goal of this thesis is to clarify whether polymorphisms have *cis*-acting regulatory properties *in vivo* on genes relevant to common infections. Given the ubiquity of RSV infection in European populations and the robust nature of the disease severity association between *IL8* and RSV-induced bronchiolitis I reasoned that investigation of polymorphisms close to the *IL8* locus would serve as a reasonable model for this investigation. My challenge was to define whether the polymorphisms studied had a functional effect on *IL8* regulation or whether these polymorphisms were merely linked to another genetic factor close to the *IL8* gene.

Prior to undertaking detailed functional studies of the haplotypes associated with RSV-induced bronchiolitis disease susceptibility I set out to develop an *in vitro* model for RSV infection. Given the importance of the respiratory epithelium to the innate immune response (outlined above) I began by establishing an *in vitro* culture of the A549 respiratory epithelial cells. Then I went on to optimize a method for purifying RSV within our laboratory. Using these resources I was able to study the kinetics of IL-8 induction in A549 cells following activation with purified RSV (pRSV), unpurified RSV (upRSV) and recombinant TNF.

In order to undertake functional studies the haplotypes described above were redefined as haplotype 1 and haplotype 2 and haplotype 3 (Figure 1.6). The -1722insT SNP was not included in this new haplotype definition as it occurs at low frequency in European

populations (0.01) and is unchanged between the two major haplotypes (Table 1.1). The unusual population structure of *IL8* in the European population allows us to examine haplotypes 1 and 2 for haplotype-specific differences in transcription using primary cells heterozygous for these haplotypes. This approach ensures that the DNA was investigated in the presence of chromatin and in the context of the usual flanking sequences that may contain important co-regulatory elements. In these circumstances important DNA-to-protein and protein-to-protein interactions that may affect gene regulation would be examined. However, as the European population is dominated by two haplotypes the results of haplotype-specific expression allow a limited insight as to which SNPs, alone or in combination, are functionally related *in vivo* to *IL8* and bronchiolitis disease severity. Therefore I sought to examine the individual SNPs in the 5' region prior to *IL8* and in the *IL8* gene with *in vitro* methods of plasmid reporter gene assays and electrophoretic mobility shift assays (EMSAs) to assess what effect the SNPs had on transcriptional regulation and protein-DNA interactions. A limitation of plasmid reporter assay systems is that only SNPs that are situated prior to the translational start site of the gene can be assessed in the context of the normal sequence of regulatory elements found in the promoter. To overcome this technical limitation EMSAs were used to systematically screen SNPs positioned beyond the translational start site for allele-specific protein-DNA binding. Where allele-specific binding occurred an attempt to identify transcription factors that were a part of differentially binding complexes was made through antibody mediated super-shift EMSAs. Where particular aspects of *cis*-acting regulation by polymorphisms were not adequately demonstrated at the *IL8* locus then experiments using the *TNF/LTA* locus were used as proof of the underlying principle that is central to this thesis.



IL8 gene	Promoter	Intron 1		Intron 2	Intron 3	3' untranslated	Frequency
SNP	-251	+396	+781	+1238	+1633	+2767	
Haplotype 1	T	T	C	insA	C	A	0.52
Haplotype 2	A	G	T	delA	T	T	0.41
Haplotype 3	A	G	C	delA	T	T	0.005

Figure 1.6

DNA sequence polymorphisms at the *IL8* locus. Schematic representation of the *IL8* promoter and gene depicting the exons (filled boxes), introns (line), the position of the polymorphisms (arrow), the different haplotypes and the haplotype frequency in a European population.

1.6.3

Concluding Points

- The severity of RSV-induced bronchiolitis is highly variable with only 5% of infants requiring hospital admission.
- RSV strain type does not appear to explain this variation in disease severity. Differences in the host immune response may prove to be a better explanation.
- Recent evidence suggests that the innate immune system and chemokines in particular play a central role in the initial host response to RSV.
- Since pulmonary polymorpholeucocyte chemotaxis is characteristic of RSV-induced bronchiolitis *IL8* is implicated as an important candidate gene as it is the principal leucocyte chemoattractant.
- Genetic epidemiology studies using both family based and case control methods show that SNPs close to and in the *IL8* gene are associated with RSV-induced bronchiolitis disease severity.
- We proposed to use *IL8* as a model system to explore the hypothesis that SNPs acting alone or in concert as haplotypes might modulate gene regulation.

CHAPTER TWO

MATERIALS AND METHODS

Materials

2.1.1

Chemical Reagents

All chemicals used were of analytical grade with the majority being obtained from Sigma-Aldrich (Poole, United Kingdom). Cell culture materials, media and additives were all of tissue culture grade and were obtained from Invitrogen (Paisley, United Kingdom (UK)) and Sigma-Aldrich (Poole, UK). Rapid-Hyb buffer, poly(dA-dT) and poly(dI-dC) were supplied by Amersham Pharmacia Biotech (Little Chalfont, UK). Luciferase assay reagent was obtained from Promega Corporation (Southampton, UK). Fugene-6 transfection reagent was obtained from Roche Diagnostics Limited (Lewes, UK). Sterile de-ionised water was used in the preparation of all solutions.

2.1.2

Cytokines, antibodies and Jurkat cell lysate

Monoclonal anti-human IL-8 antibody was obtained from R and D systems (Abingdon, UK) whilst specific rabbit polyclonal antiserum to respiratory syncytial virus was purchased from Biogenesis (Poole, UK). Recombinant human interleukin-1β and recombinant human interferon-γ were purchased from Roche Diagnostics Limited. 4α-Phorbol 12-myristate 13-acetate (PMA) and TNF and ionomycin was obtained from Sigma-Aldrich (Poole, UK). Anti-c-Myb (clone 1-1) antibody was obtained from Upstate biotechnology (Lake Placid, New York, United States of America (USA)). Anti-c-Myb (C-19 and M19) antibodies, anti-

B-Myb (C-20) antibodies, anti-Oct-1 antibodies, anti-NFκB antibodies and anti-C/EBP β (C-19) antibodies were obtained from Santa Cruz Biotechnology (Santa Cruz, USA). Jurkat cell lysate was obtained from Upstate biotechnology (Lake Placid, New York, USA). Anti-mouse Immunoglobulin horseradish peroxidase linked whole antibody from sheep was obtained from Amersham Pharmacia Biotech. Antibodies against phosphylated serine residues of the C-terminal domain (CTD) of Pol II (Ser5, MMS-134R clone H14; Ser2, MMS-129R clone H5) were obtained from Covance (California, USA). Antibodies to anti-T antigen pAb101 (sc-147) were obtained from Santa Cruz Biotechnology (Santa Cruz, USA).

2.1.3

Enzymes

Restriction endonucleases *Bgl* II and *Nhe* I were obtained from Roche Diagnostics (Lewes, United Kingdom). Restriction endonuclease *Bst* Z 171 was obtained from New England Biolabs (Hitchin, UK). Biotaq DNA polymerase was obtained from Bioline (London, UK).

2.1.4

Northern and Western blots

All Northern blots and Western blots were performed using Hybond-N+ membranes and Hybond-P membranes respectively (Amersham Pharmacia Biotech). Western blots were run on the Xcell II blot module using NOVEX pre cast NuPAGE Bis-Tris SDS-PAGE mini-gels all obtained from Invitrogen (Paisley, UK).

2.1.5

Xray film, Luminometer and Phosphoimager equipment

All autoradiography was performed using X-OMAT AR film from Kodak (Hemel Hempstead, United Kingdom). Cyclone storage phosphor screen and imager were obtained from Packard Bioscience (Pangbourne, UK). Luminometer readings were obtained from a Turner TD-20e Luminometer (Sunnyvale, California, USA).

2.1.6

Molecular Kits

Promega Corporation	Dual-Luciferase reporter assay system
Qiagen Ltd (Crawley, UK)	Qiagen Plasmid Midi Kit
Amersham Pharmacia Biotech	Megaprime DNA label system
Amersham Pharmacia Biotech	QuickPrep *micro* mRNA Purification
Amersham Pharmacia Biotech	First-Strand cDNA synthesis Kit
Amersham Pharmacia Biotech	ECL+Plus Western blotting detection system

2.1.7

Nucleotides

Nucleotides [α-^{32}P] dCTP, 371.2 MBq/ml were supplied by Amersham Pharmacia Biotech.

2.1.8

RNA and DNA Size markers

RNA size markers for Northern Blot agarose gel electrophoresis were obtained from Sigma-Aldrich. DNA size markers for agarose gel electrophoresis were obtained from Invitrogen (Paisley, UK).

2.1.9

DNA Primers

All DNA primers and oligonucleotide probes were obtained from Invitrogen (Paisley, UK) and MWG-Biotech (Ebersberg, Germany).

2.1.10

Virus and Human Cell Lines

Human HEp-2 laryngeal epithelial cells, Human A549 pulmonary type II epithelial cells, Human H460 pulmonary epithelial cells, acute T-cell leukaemia cell line Jurkat and human long strain of RSV (A2) and COS-7 cells were obtained from the American Type Culture Collection, (Manassas, Va.). EBV-transformed B cell lines were obtained from the Centre d'Etude du Polymorphisme Humain (CEPH) collection (Coriell Institute for Medical Research). Competent DH5α bacterial cells of library efficiency were obtained from Invitrogen (Paisley, UK).

Methods

2.2.1

A549 cell culture

A549 cells were cultured in Kaighn's modified media ($F_{12}K$) (Invitrogen, Paisley, UK) supplemented with 100U/ml penicillin, 100 mg/ml streptomycin, 2 mM L-glutamine and 10% foetal bovine serum and incubated at 37°C in 5% CO_2.

2.2.2

H460 cell culture

H460 cells were cultured in RPMI medium supplemented with 100U/ml penicillin, 100 mg/ml streptomycin, 2 mM L-glutamine, 1.5g/l sodium bicarbonate, 4.5g/l glucose, 10mM HEPES and 10% fetal bovine serum. Cells incubated at 37°C in 5% CO_2

2.2.3

HEp2 cell culture

Human HEp-2 laryngeal epithelial cells were grown in Dulbecco's modified eagles medium (DMEM) (Sigma-Aldrich) supplemented with 100U/ml penicillin, 100 mg/ml streptomycin, 2 mM L-glutamine and 5% foetal bovine serum. Cells were incubated at 37°C in 5% CO_2.

2.2.4

Jurkat cell culture

Jurkat cells were cultured in RPMI medium supplemented with 100U/ml penicillin, 100 mg/ml streptomycin, 2 mM L-glutamine and 10% foetal bovine serum. Cells were incubated at 37°C in 5% CO_2

2.2.5

COS cell culture

COS-7 cells were cultured in DMEM supplemented with 10% foetal bovine serum, 100U of penicillin per ml, 100mg of streptomycin per ml, 0.2mM L-glutamine and 0.1% glucose. Cells were incubated at 37°C in 5% CO_2 Cell culture of COS-7 cells was performed by Senior Research Fellow Irina Udalova and used with her permission.

2.2.6

EBV-transformed B cell line culture

EBV-transformed B cell lines from the CEPH collection (Coriell Institute for Medical Research) were cultured to a cell count of 16 x 10^8 cells per ml in RPMI medium supplemented with 100U/ml penicillin, 100 mg/ml streptomycin, 2 mM L-glutamine and 10% foetal bovine serum at 37°C in 5% CO_2 Cells were stimulated with 200nM PMA and 1 μM ionomycin. EBV-transformed B cell line culture and stimulation was performed by Brendan Keating, D.Phil student.

2.2.7

Primary respiratory epithelial cell harvest and culture

Consent to approach patients for anonymous tissue collection was obtained from the Oxfordshire research ethics committee (OXREC) number 99/5/27. Healthy adult patients admitted to the John Radcliffe Hospitals NHS Trust for routine day surgery were consented for primary nasal epithelial cell harvest by oribrush cell collector (Orifice Medical AB, Henleys Medical, Welwyn Garden City, UK) brushes under general anaesthetic. Cell

collections were taken by brushing on the lateral wall of the pharynx below the inferior turbinate. The cells from each nostril were then pooled in 4 mls of Kaighn's modified media ($F_{12}K$) supplemented with 100U/ml penicillin, 100 mg/ml streptomycin, 2 mM L-glutamine and 10% foetal bovine serum and 30 ng/ml TNF. They were then divided into two aliquots of 2 mls and incubated at 37°C in 5% CO_2 for three and six hours respectively. All patients had a 5 ml venous blood sample taken for genotyping for -251 and +2767 SNPs and to act as a source of genomic DNA. Regular smokers were excluded as they were likely to have a significant proportion of stratified squamous epithelia present in their pharynx as opposed to the normal pseudo-stratified ciliated columnar epithelium. All cell collections were checked under a microscope for the presence of motile cilia indicative of the presence of ciliated columnar epithelial cells. In total 30 patients were recruited all of whom were Caucasian in ethnic origin. There were 21 women and 9 men.

2.2.8

RSV culture

HEp-2 cells were cultured in DMEM supplemented with with 100U/ml penicillin, 100 mg/ml streptomycin, 2 mM L-glutamine, 2.5% foetal bovine serum and 20mM Hepes. When confluent the cells were infected with unpurified RSV at a multiplicity of infection (MOI) of 1. After two hours of incubation at 37 °C the media were replaced with fresh medium. The cells were incubated for a further 48 hours until viral cytopathological changes were present in 75 to 100% of the HEp-2 cells. The RSV virus was then harvested according to the purification protocol detailed below. For the experiments, A549 cells were grown to approximately a 60% monolayer in $F_{12}K$ media supplemented with 100U/ml penicillin, 100 mg/ml streptomycin, 2 mM L-glutamine 2.5% fetal bovine serum and 20mM

Hepes. During this logarithmic growth phase cells were infected with purified RSV at a multiplicity of infection (MOI) of 1. After two hours of incubation at 37 °C the media were replaced with fresh medium.

2.2.9

RSV purification

RSV purification was performed using a protocol modified from a previously published method (Ueba 1978). Infected HEp-2 cells were scraped from the flask with glass beads after which the supernatant was collected and centrifuged at 2,000g for 15 minutes at 4°C. The RSV supernatant was then pooled and mixed for 2 hours at 4°C with 50% polyethylene glycol (PEG) such that the final concentration was 10% PEG. Then the RSV supernatant/PEG mix was centrifuged at 10,000g for 30 minutes at 4°C. The resulting supernatant was divided and placed on a discontinuous sucrose gradient of 30% sucrose and 60% sucrose diluted in $F_{12}K$ media supplemented with 100U/ml penicillin, 100 mg/ml streptomycin, 2 mM L-glutamine 2.5% foetal bovine serum and 20mM Hepes. This gradient was then centrifuged at 82,000g at 4°C for 90 minutes. A white band containing the purified RSV was identified within on the sucrose gradient, isolated and pooled. The pooled concentrated RSV was then divided into aliquots and snap frozen in dry ice and ethanol after which it was stored for 24 hours at –80 °C and thereafter liquid nitrogen. No TNF was found in the sucrose purified RSV (pRSV) preparation.

2.2.10

RSV inactivation

RSV was inactivated following exposure to a 254 nm Ultra-violet light source for a time of three minutes.

2.2.11

RSV protein staining

To confirm that the cells were infected with RSV, cultures were stained for viral protein 24 hours post-infection using a specific rabbit polyclonal antiserum.

2.2.12

Northern Blot

Six well plates containing the cultured A549 cells were placed on ice after which the cells were washed twice with pre cooled phosphate buffered saline (PBS) at 4^0C. TRI reagent (Sigma-Aldrich) was then added and mixed to achieve a homogeneous lysate. RNA extraction then proceeded as per protocol (Sigma-Aldrich 1999). Extracted RNA was heated at 65^0C for 10 minutes after which it was cooled on ice and run on a denaturing 1.4% agarose and formaldehyde gel. Movement of the RNA to the Hybond-N+ membrane was performed by overnight capillary transfer. The RNA was subsequently stabilized with UV-cross linking. *IL8* and β-actin DNA probes were radio-labeled with [α-^{32}P] dATP using the Megaprime DNA label system and after heat denaturation at 98^0C were hybridised to the membrane in the presence of Rapid-Hyb buffer at 65^0C overnight. Unhybridised probe was removed by stringency washing in SSC and SDS buffers (Sigma-Aldrich) before the membrane was exposed to OMAT AR film (Kodak).

2.2.13

Nitric Oxide assay

In these experiments reactive nitrogen intermediates (RNI) were measured as these include all the immediate products of the NOS reaction (.NO radical, NO^-, NO^+) and their resulting conversion products such as NO2, NO2-, NO3, and peroxynitrite (ONOO). Reactive nitrite intermediates (RNI) were measured in culture supernatants using the Griess reaction after reduction of nitrate to nitrite as previously described (Rockett et al. 1998). N^G-monomethyl-L-arginine (L-NMMA) was synthesized as previously described (Patthy et al. 1977).

2.2.14

Total Protein Assay

A549 cells were washed twice with PBS and lysed with passive lysis buffer preparation (Promega Corporation). Protein contents of this cellular extract were measured by a spectrophotometric (Bradford) assay using Bovine serum albumin as a standard (Bio-Rad protein detection reagent, Bio-Rad Laboratories, Hemel Hempstead, UK).

2.2.15

Enzyme linked immunosorbant assay (ELISA)

IFNγ, TNF and IL-1β were quantitated in cell culture supernatants by double antibody (sandwich) quantitative enzyme-linked immunosorbent assay (ELISA) (Smith 1993) kits (R&D Systems, Abingdon, United Kingdom) according to manufacturer's protocol. The sensitivity of detection of the IFNγ, TNF, IL-1β and IL-8 assays were 15.6 pg/ml, 15.6 pg/ml, 3.9 pg/ml and 10 pg/ml respectively.

2.2.16

IL8 construct

The 1.4kb 5'-flanking promoter region of the *IL8* gene was obtained by polymerase chain reaction (PCR) *IL8* gene amplification of human genomic DNA and cloned into a basic pGL2 vector that contains the firefly luciferase gene in the absence of DNA that confers any transcriptional activity. In order to examine the relative importance of this gene region to IL-8 expression truncated constructs at –1019, -789, -389, -169, -109 and -49 base pairs from the transcription start site were also constructed. Construct design and synthesis was performed by Senior Research Fellow Irina Udalova and used with her permission.

2.2.17

C/EBP β construct

The original CMV-NFIL6 construct was a kind gift from Dr Akira. The EcoRI/SalI fragment, encoding the full length protein was subsequently recloned into EcoRI/XhoI pcDNA3.

2.2.18

IL8 construct bacterial transformation, culture and purification

Competent DH5α bacterial cells (25 μλ) were transformed with the interleukin-8 construct by heat shock at 37 ^{0}C for 45 seconds followed by immediate incubation for 2 minutes at 4^{0}C. Following selective culture on LB/ampicillin plates (50 μg/ml) a single colony was grown up in liquid LB media overnight after which the IL-8 construct was extracted using

the Qiagen Plasmid Midi Kit according to the manufacture's instructions. Construct integrity was checked by restriction endonuclease digest with *Bgl* II and *Nhe* I enzymes.

2.2.19

Interleukin-8 construct transfection into A549 and H460 cells

A549 cells were transfected using the FuGene 6 lipid-based transfection reagent (Roche Diagnostics Limited). On the day before transfection A549 and H460 cells were seeded at 60,000 cells/cm^2 to achieve 50 to 80% confluence 16 hours later. Fugene 3 µl per well was placed in 100 µl of $F_{12}K$ after which 1 µg of plasmid DNA solution (0.5 µg/µl) was added. After incubation for 45 minutes at room temperature 100 µl transfection mix was added to each 35 mm well which contained 2ml of $F_{12}K$ serum containing media. After 24 hours the cells were stimulated with either RSV or cytokines and then harvested at the appropriate time point according to the Promega luciferase assay system protocol. Briefly, A549 cells were placed on ice after which the cells were washed twice with PBS at 4^0C. They were then lysed with 250 µl of Promega cell culture lysis reagent, scrapped from the well and stored at -70^0C until assayed. Prior to the assay all samples were centrifuged to remove cell debris. Luciferase activity measurement was made by mixing 10 µl of cell lysate with 90 µl of Promega Luciferase reagent II. Peak light intensity was determined on a luminometer. Luciferase measurements were corrected for protein content of the cell lysate, by Total Protein assay protocol as detailed below, and expressed in terms of fold activation of stimulated over unstimulated cell lysate luciferase values. To control for transfection efficiency a pRL-TK plasmid co-reporter vector was co-transfected with the experimental vector at a ratio of 1:50. The pRL-TK contains a herpes simplex virus thymidine kinase

promoter upstream of a renilla luciferase gene. Assay of the luciferase signal and correction with the renilla luminescence was performed using the Dual-Luciferase reporter assay system.

2.2.20

Nuclear extracts and Electrophoretic mobility shift assay (EMSA)

Oligonucleotide probes were annealed at 100 degrees and slowly cooled to room temperature after whch they were radiolabeled with [α-^{32}P]dCTP using Klenow enzyme (Roche). The individual probes used were:

251A	(F: agctGAAATAAAAAAGCATACAATTGATAATTCACCAAATTG; R: agctCAATTTGGTGAATTATCAATTGTATGCTTTTTTATTTC);
251T	(F: agctGAAATAAAAAAGCATACATTTGATAATTCACCAAATTG; R: agctCAATTTGGTGAATTATCAAATGTATGCTTTTTTATTTC);
396G	(F: agctATTTATACCAGGTAGCATGCA; R: agctTGCATGCTACCTGGTATAAAT);
396T	(F: agctATTTATACCATGTAGCATGCA; R: agctTGCATGCTACATGGTATAAAT);
781T	(F: agctTATAGGAAGTTGTTCAATGTT; R: agctAACATTGAACAACTTCCTATA);
781C	(F: agctTATAGGAAGTCGTTCAATGTT; R: agctAACATTGAACGACTTCCTATA);
1238delA	(F: agctCAAAACAAAATAAAAATATTT; R: agctAAATATTTTTATTTTGTTTTG);
1238insA	(F: agctCAAAACAAAAATAAAAATATTT; R: agctAAATATTTTTATTTTTGTTTTG);
1633T	(F: agctTTTGTTGTACTTATGACCAGAA; R: agctTTCTGGTCATAAGTACAACAA);
1633C	(F: agctTTTGTTGTACTCATGACCAGAA; R: agctTTCTGGTCATGAGTACAACAA);
2767T	(F: agctATTTCAGATATACAACAAATA; R: agctTATTTGTTGTATATGTGAAAT);
2767A	(F: agctATTTCAGATAAACAACAAATA; R: agctTATTTGTTGTTTATGTGAAAT)
c-Myb wild type	(F: agctTCTACACCCTAACTGACACAT; R: agctATGTGTCAGTTAGGGTGTAGA)
c-Myb mutant	(F: agctTCTACACCCTGGCTGACACAT; R: agctATGTGTCAGCCAGGGTGTAGA)

A549 cells were stimulated with pRSV at an MOI of 1 and harvested at 0, 2, 4, 8, 16, 24 and 48 hours after activation. Jurkat cells were stimulated with PMA (50 ng/ml) for a duration of 20 hours. Nuclear extracts were prepared as described previously (Schreiber et al. 1989). The binding reaction contained 12mM HEPES, pH 7.8, 80-100 mM KCl, 1 mM EDTA, 1 mM EGTA, 12% glycerol, and 0.125 to 0.25 μg/ml of poly(dI-dC) or poly(dA-dT). Nuclear extracts (2 μl) were mixed in an 8 μl reaction with 0.2-0.5 ng of labeled probe (1 to 5 x 10^4 counts per minute) and incubated at room temperature for 10 minutes. The reaction was analysed by electrophoresis in a nondenaturing 5% polyacrylamide gel at 4^0C in 0.5xTBE buffer. For supershift analysis, the reaction mixture was incubated with appropriate antiserum at room temperature for 5-10 minutes prior to the addition of the labeled probe. Indicated gels were quantified using the Phosphoimager.

2.2.21

UV Cross linking

The binding reaction was performed with radiolabelled oligoduplex corresponding to the -251A in which a dT close to the SNP were substituted with bromodeoxyuridine (BrdU) as follows:

251.BrdUA (F: agctGAAATAAAAAAGCATACAAT(BrdU)GATAATTCACCAAATTG;
R: agctCAATTTGGTGAATTATCAATTGTATGCTTTTTTATTTC);.

The EMSA gel was UV illuminated at 302nm for 30 minutes at 4^0C and exposed to autoradiography for 2 to 4 hours at the same temperature. The region corresponding to the DNA-protein complex was excised and the proteins were eluted in 2xSDS buffer (100mM Tris-Cl, ph 6.8, 200mM dithiothreitol, 4% SDS, 20% glycerol) at 37^0C overnight. Final

products were analysed on either a 10% Bis-Tris MOPS-SDS NuPAGE gel (Invitrogen) or on a 4-12% Bis Tris MOPS-SDS NuPAGE gel (Invitrogen).

2.2.22

mRNA extraction and cDNA synthesis

For IL-8 transcripts mRNA extraction and cDNA synthesis were performed with QuickPrep *micro* mRNA Purification kit and First-Strand cDNA synthesis kit respectively according to manufacturer's instructions. For transcripts from EBV-transformed B cell lines total RNA was isolated using TRI reagent (Sigma). Following this cDNA was prepared with random decamers and AMV reverse transcriptase after which relative quantitative RT-PCR (Ambion) was performed according to manufacturer's instructions. Transcripts from EBV-transformed cell lines were prepared by Brendan Keating, D.Phil student.

2.2.23

Semiquantitative Polymerase Chain Reaction and allelic-specific restriction digest otherwise known as mutant analysis by PCR and restriction enzyme cleavage (MAPREC)

Genomic DNA and cDNA dissolved in TE were assayed by semiquantitative PCR as described previously (Takahashi et al. 2000) with the following primers:

[1] F: ACCCAGTTAAATTTTCATTTCAGGTA; [2] R: CAGCTGGCAATGACAAGACTG.

First round PCR amplification was performed using 5ng genomic DNA or cDNA in a 15 µl reaction volume using 0.25U Biotaq in 1x PCR reaction buffer with 0.4 mM dNTPs, 1.9 mM $MgCl_2$, and 0.2 µM each primer. The reactions were denatured at 94^{0}C for 2 minutes followed by 28 cycles of 94^{0}C for 30 seconds, 60^{0}C for 30 seconds, 72^{0}C for 30 seconds

followed by final extension at 72°C for 10 minutes. Amplification was carried out in skirted 96 well plates using an MJ Tetrad thermal cycler. To ensure the PCR reaction produced adequate amplification a 5µl aliquot from each reaction was then electrophoresed on a 5% agarose gel. A 1.5µl volume PCR reaction mix containing 1x PCR reaction buffer with 0.4 mM dNTPs, 1.9 mM $MgCl_2$, and 0.02 mM [α-^{32}P]dCTP was then added to each reaction and denatured at 94°C for 2 minutes followed by 60°C for 30 seconds and 72°C for 15 minutes. Primer [1] was designed to introduce a *Bst* Z 171 restriction site in the 2767T allele. The amplification products were restriction digested with *Bst* Z 171 and then electrophoresed on an 8% non-denaturing polyacrylamide gel. The quantity of amplification product was then quantified using the Phosphoimager.

2.2.24

Primer extension and Mass spectrometry (PE/MS) for IL8 transcript allele-specific quantitation

First round PCR amplification was performed using 50ng genomic DNA or cDNA in a 20 µl reaction volume using 0.025U Biotaq in 1x PCR reaction buffer with 0.8mM dNTPs, 1.9mM $MgCl_2$, and 0.2µM each primer ([1] F: ACGTTGGATGTCAGATAAACAATAATGTAC; [2] R: ACGTTGGATGGGGTACCCAGTTAAATTTTC). The reactions were denatured at 96°C for 4 minutes followed by 5 cycles of 96°C for 45 seconds, 56°C for 45 seconds, 72°C for 30 seconds and then 30 cycles of 94°C for 45 seconds, 65°C for 45 seconds, 72°C for 30 seconds and finished by a final extension at 72°C for 10 minutes. Amplification was carried out in skirted 96 well plates using an MJ Tetrad thermal cycler. Each PCR reaction was

subaliquoted into 4 wells on a 384 well plate, 4µl per well. Non-incorporated dNTPs were removed using shrimp alkaline phosphatase by incubating at 37^{0}C for 20 minutes followed by 85^{0}C for 5 minutes. Primer extension was performed using homogeneous MassEXTEND (Sequenom) reaction comprising a mix of 100 µM extension primer (CCAGTTAAATTTTCATTTCAGATA), 0.576U MassEXTEND enzyme, buffer and a termination mix of dATP, ddTTP, ddCTP and ddGTP. Primer extension was performed at 94^{0}C for 2 minutes then 40 cycles of 94^{0}C for 5 seconds, 52^{0}C for 5 seconds and 72^{0}C for 5 seconds. Products of primer extension were desalted using SpectroCLEAN (Sequenom) resin and transferred onto a SpectroCHIP (Sequenom) microarray by SpectroPOINT (Sequenom) nanolitre dispenser. Matrix-assisted laser desorption time-of-flight mass spectrometry (MALDI-TOF) analysis was performed using a SpectroREADER (Sequenom) mass spectrometer.

2.2.25

Preparation of Chromatin

Cells were lysed and sonicated using a protocol modified from that which has been previously described (Takahashi et al. 2000). Approximately, 16 x 10^8 cells were cross-linked using formaldehyde at a final concentration of 1% for 45 minutes at room temperature after which the reaction was quenched using glycine at a final concentration of 0.125M. Cells were lysed and the nuclei collected by incubating and then centrifuging the cell pellets in the following buffers: buffer A 20mls (50mM Hepes-KOH pH 7.5, 140mM NaCl, 1mM EDTA, 10% Glycerol, 0.5% NP-40, 0.25% Triton X-100, 1x Complete protease inhibitors (Roche) 1mM benzamidine, 50µg/ml TLCK, 50µg/ml TPCK, 1µg/ml pepstatin, 10mM sodium

pyrophosphate) at 4°C for 10 minutes; buffer B 16 mls (200mM NaCl, 1mM EDTA pH 8, 0.5mM EGTA pH 8, 10mM Tris pH 8, 1x Complete protease inhibitors (Roche) 1mM benzamidine, 50μg/ml TLCK, 50μg/ml TPCK, 1μg/ml pepstatin, 10mM sodium pyrophosphate) at 25°C for 10 minutes and finally buffer C 10mls (1mM EDTA pH 8, 0.5mM EGTA pH 8, 10mM Tris-HCl pH 8, 1x Complete protease inhibitors (Roche) 1mM benzamidine, 50μg/ml TLCK, 50μg/ml TPCK, 1μg/ml pepstatin, 10mM sodium pyrophosphate) at 4°C for 10 minutes. Cells were sonicated at 4°C using a Branson 450 sonifier using a medium tip first for 40 cycles (4 second burst with 8 seconds pause) at an amplitude of 20% followed by 40 cycles (4 second burst with 8 seconds pause) at an amplitude of 30%. The lysed sonicated cells were then centrifuged, the pellets discarded and the chromatin supernatant saved for further processing.

In experiments to optimize the purification of chromatin fragments chromatin suspension was adjusted to contain 0.5% N-laurosarcosine and placed on a discontinuous cesium chloride (CsCl) gradient as follows: 30.29g, 20.21g and 12.07g of CsCl were diluted in 30 ml of Tris-EDTA (TE) solution (10mM Tris-HCl pH8, 1mM EDTA) to give 1.75g/cm^3, 1.5g/cm^3 and 1.3g/cm^3 density CsCl solutions respectively. CsCl gradients are formed by gently overlaying successively less dense solutions in ultraclear centrifuge tubes 14 x 89mm (Beckman). Chromatin solutions were added to the top of the CsCl gradient and centrifuged at 165,000g for 20 hours at 20%. After centrifugation 1ml fractions were removed from the gradient and the distribution of chromatin estimated by running a 12μl aliquot of each on a 1% TBE agarose gel. Fractions containing chromatin were pooled (usually 3 to 5 in

number) and dialysed overnight in 2 litres of dialysis buffer (10mM Tris-HCl pH 8, 1mM EDTA pH 8, 0.5mM EGTA pH 8, 10% glycerol, 10mM sodium pyrophosphate) using spectra/por (Fisher) dialysis membrane. The Chromatin was divided into 1ml aliquots, snap frozen in dry ice and ethanol and then stored at -80°C

When not ustilising caesium chloride gradient purification a small aliquot of the Chromatin supernatant was diluted with N-laurosarcosine (Sigma) to 0.5% and run on a 1.4% TBE agarose gel to ensure that the chromatin fragments were sonicated to between 0.5 and 3 kilobases. A second small aliquot was used to estimate the chromatin concentration by OD quantitation. The Chromatin was then diluted with glycerol to a final concentration of 10%, divided into 1ml aliquots, snap frozen in dry ice and ethanol and then stored at -80°C.

2.2.26

Chromatin immunoprecipitation

Immunoprecipitation of chromatin was achieved through the use of magnetic beads. Dynabeads M-280 precoated with sheep anti-mouse IgG (Dynal) were washed twice in Phosphate buffered saline containing 5mg/ml BSA (PBS-BSA) (Sigma) and then incubated overnight at 4°C on a nutator with the relevant mouse monoclonal antibodies in PBS-BSA. For the immunoprecipitation of phosphorylated Pol II specific antibodies against phosphorylated serine residues of the C-terminal domain (CTD) of the enzyme (Ser5, MMS-134R clone H14; Ser2, MMS-129R clone H5; Covance, Princeton, USA) were used. As a control a mock antibody raised against the anti-T antigen pAb101 (sc-147; Santa Cruz Biotechnology, Santa Cruz, USA) was used. Unbound antibody was removed from the

beads by washing three times with PBS-BSA after which the antibody-beads were incubated overnight at 4°C on a nutator with 200µg chromatin in 1x RIPA buffer (10mM Tris HCl pH 8, 1mM EDTA, 0.5mM EGTA, 1% Triton X-100, 0.1% sodium deoxycholate, 0.1% SDS, 560mM NaCl) containing in addition 5mM Sodium pyrophosphate, 1x complete protease inhibitors (Roche) and 5µg/ml pepstatin (Roche). The bead-chromatin immunoprecipitations were washed twice with 1x RIPA-POL buffer (10mM Tris HCl pH 8, 1mM EDTA, 0.5mM EGTA, 1% Triton X-100, 0.1% sodium deoxycholate, 1% sodium deoxycholate, 0.1% SDS, 1x complete protease inhibitors, 5µg/ml pepstatin 10mM sodium pyrophosphate) containing 140mM NaCl, once with 1x RIPA-POL buffer containing 100µg/ml Salmon sperm, once with 1x RIPA-POL buffer containing 100µg/ml Salmon sperm and 300mM NaCl and once with 1x RIPA-POL buffer containing 250mM LiCl. To remove the monoclonal antibody the washed beads were then re-suspended beads in 50µl of TE digestion buffer containing 1% SDS and incubated at 65°C for 10 minutes. The beads were pelleted by centrifugation after which 120µl of TE was added. In order to reverse the crosslinks between the transcription factors and the DNA the beads were incubated at 65°C overnight to reverse cross-links. The proteins that were associated with the DNA were then digested by the addition of proteinase K to a final concentration of 200µg/ml and incubated at 37°C for 2 hours. The DNA was then purified with a phenol-chloroform extraction and precipitated with ethanol in the presence of 20µg glycogen. To remove RNA from the preparation the pellets were re-suspended in 30µl TE containing 10µg RNase A and incubated at 37°C for 1 hour. Finally the DNA was purified through a spin column (Qiagen) and eluted in 200µl of TE. In order to ensure the specificity of each immunoprecipitation a sample of DNA template derived from Anti-H14, Anti-H5 and

Anti-SV40 antibody immunoprecipitations was amplified under the following conditions: denatured at 96^{0}C for 1 minutes followed by 5 cycles of 96^{0}C for 35 seconds, 68^{0}C for 45 seconds, 72^{0}C for 35 seconds and then 21 cycles of 96^{0}C for 25 seconds, 65^{0}C for 50 seconds, 72^{0}C for 40 seconds and finished by 6 cycles of 96^{0}C for 35 seconds, 56^{0}C for 60 seconds, 72^{0}C for 1 minute 30 seconds. The products were run on a 1% TBE agarose gel. No significant DNA was detected after immunoprecipitation using anti-SV40 antibodies.

2.2.27

Primer extension and Mass spectrometry (PE/MS) for TNF/LTA locus allele-specific quantitation by haploCHIP

First round PCR amplification was performed using 50ng genomic DNA or cDNA in a 10µl reaction volume using 0.2U Biotaq in 1x PCR reaction buffer with 0.2mM dNTPs, 1.9mM $MgCl_2$, and 0.2µM each primer. The following primers were used:

LTA+10 (1st: ACGTTGGATGTATAAAGGGACCTGAGCGTC;
2nd: ACGTTGGATGTAGTCCAAAGCACGAAGCAC);

LTA+80 (1st: ACGTTGGATGATCCAGGCAGCAGGTGCAG;
2nd: ACGTTGGATGCCGTGCTTCGTGCTTTGGAC);

LTA+252 (1st: ACGTTGGATGGAGAGACAGGAAGGGAACAG;
2nd: ACGTTGGATGACTCTCCATCTGTCAGTCTC);

LTA+368 (1st ACGTTGGATGCTGCATCTTGTCCCCTTCTC
2nd: ACGTTGGATGAAGACAGACCTCCCGCCCT);

LTA+723 (1st: ACGTTGGATGTCAGCACCCCAAGATGCATC;
2nd: ACGTTGGATGGAGGTCAGGTGGATGTTTAC);

TNF-1032 (1st: ACGTTGGATGGGGAAGCAAAGGAGAAGCTG;
2nd: ACGTTGGATGTACATGTGGCCATATCTCCC);

The reactions were denatured at 96°C for 1 minute followed by 5 cycles of 96°C for 35 seconds, 68°C for 45 seconds, 72°C for 35 seconds and then 21 cycles of 96°C for 25 seconds, 65°C for 50 seconds, 72°C for 40 seconds and finished by 6 cycles of 96°C for 35 seconds, 56°C for 60 seconds, 72°C for 1 minute 30 seconds. Amplification was carried out in skirted 96 well plates using an MJ Tetrad thermal cycler. Each PCR reaction was sub-aliquoted into 2 wells on a 384 well plate, 4μl per well. Non-incorporated dNTPs were removed using shrimp alkaline phosphatase by incubating at 37°C for 20 minutes followed by 85°C for 5 minutes. Primer extension was performed using homogeneous MassEXTEND (Sequenom) reaction comprising a mix of 100 μM extension primer, 0.576U MassEXTEND enzyme, buffer and a termination mix of ddATP, ddTTP, ddCTP and dGTP. The following extension primers were used:

LTA+10	GAGAGCCTCACCTGCTGTG
LTA+80	GTGCTTTGGACTACCGCCC
LTA+252	CACATTCTCTGTTTCTGCCATG
LTA+368	CCTCCCGCCCTGGGAGA
LTA+723	GGTGAGCAGCAGGTTTGAGG
TNF-1032	CCAGACCCTGACTTTTCCTTC

Primer extension was performed at 94°C for 2 minutes then 40 cycles of 94°C for 5 seconds, 52°C for 5 seconds and 72°C for 5 seconds. Products of primer extension were desalted using SpectroCLEAN (Sequenom) resin and transferred onto a SpectroCHIP (Sequenom) microarray by SpectroPOINT (Sequenom) nanolitre dispenser. Matrix-assisted laser desorption time-of-flight mass spectrometry (MALDI-TOF) analysis was performed using a SpectroREADER (Sequenom) mass spectrometer. In order to control for random variation

between experiments immunoprecipitated DNA was paired with genomic DNA from the same cell line on each chip and then assayed by PE/MS.

2.2.28

Western Blotting

A549 and Jurkat nuclear extract preparations as well as Jurkat cell lysate preparations were denatured by heating at 75^{0}C for 5 minutes and then run on a NuPAGE Bis-Tris SDS-PAGE pre-cast mini-gel using NuPAGE MEPS SDS running buffer after which the gels were blotted onto Hybond-P membranes at 100 V over 1 hour. The membranes were then dried as directed by the manufacturer's instructions (Invitrogen, Paisley, UK). c-Myb and C/EBP β proteins were bound with rabbit polyclonal M-19 anti-c-Myb antibody (Santa Cruz) and anti-C/EBP β (C-19) antibodies (Santa Cruz) respectively. Detection of bound antibody occurred with streptavidin-conjugated anti-rabbit antibody (Amerhsam Pharmacia) and then ECL solution according to manufacturer's instructions (Amersham Pharmacia). The western blot was exposed to OMAT AR film (Kodak) for 10 to 120 seconds.

2.2.29

Statistics

Data are summarized as means with standard errors, geometric means with 95% confidence intervals or medians with an inter-quartile range as most appropriate depending on whether they were normally distributed. The significance of differences in mean values, where the data had a normal distribution, was determined using Student's *t* test or in the case of paired data with a paired *t* test. The sd test was used in parametric analysis to check the assumption of equal variance between groups. A two-sided p value was considered significant if below

0.05. All analyses were performed using the statistical software package STATA Versions 6 and 7 (Stata Corporation, College Station, Texas, USA).

CHAPTER THREE

ESTABLISHING AN *IN VITRO* MODEL OF RSV INFECTION

3.1.1

Introduction

Early experiments of unpurified RSV (upRSV) *in vitro* infection of human bronchial epithelial cell lines showed that multiple cytokines were induced by viral infection including IL-8, IL-6 and granulocyte macrophage colony-stimulating factor (GMCSF) (Noah, 1993). These *in vitro* studies were confirmed with *in vivo* studies on children suffering from upper respiratory tract RSV infection where nasal lavage fluid showed the presence of IL-1β, IL-8, IL-6 and TNF (Noah, 1995). Focused investigation into the role of IL-8, as a chemotactic and activating cytokine, began with the observation that polymorphonuclear neutrophils (PMN) accounted for 70% of the inflammatory cells in the lower airway and 93% of the inflammatory cells in the nasal lavage samples (Everard, 1994).

The initial experiments on upRSV induced IL-8 expression in Human A549 pulmonary type II epithelial cells (A549) showed very early *IL8* transcription and a dose response relationship between the upRSV multiplicity of infection (MOI) and IL-8 which was observed for both active and inactive virus (Fiedler et al. 1995). A possible explanation for these data is that contaminating cytokines in the upRSV preparation could have been activating *IL8* expression. Transfection experiments examining the 5'-flanking region of the

IL8 gene with upRSV as a stimulus showed up to a 13 fold activation of the reporter gene (Fiedler et al. 1996).

However, subsequent studies showed that infection of A549 cells with purified RSV (pRSV) induced expression of a number of IL-8 activating cytokines such as IL-1β, TNF, IFNγ and IL-1α (Patel et al. 1995). To remove these cytokines RSV was purified with polyethylene glycol precipitation and centrifuge sucrose density gradients (Ueba, 1978). Moreover, pRSV stimulation of *IL8* promoter constructs transfected into A549 cells showed only a 4 to 5 fold activation of the reporter gene over uninfected controls (Casola et al, 2000).

Prior to embarking on *in vitro* studies on the effect polymorphisms had on haplotype-specific *IL8* transcription I sought to establish within our laboratory the purification method of RSV and establish an *in vitro* model of RSV infection of respiratory epithelial cells. I wished to clarify the pattern of upRSV and pRSV induced *IL8* expression in A549 cells and compare this to that seen with TNF induced *IL8* expression.

Results

3.2.1

Purified RSV infects and replicates in A549 cells

I sought to test whether A549 cells were infected with RSV and whether RSV was able to form syncytia in this cell line. A549 cells were infected with pRSV at an MOI of 0.5. The presence of RSV protein expression was confirmed up to 24 hours through immunocytochemical staining using polyclonal anti-RSV antibodies (Table 3.1). I observed a steady increase in RSV protein expression and syncytia formation prior to 48 hours following the infection both of which indicate active viral replication.

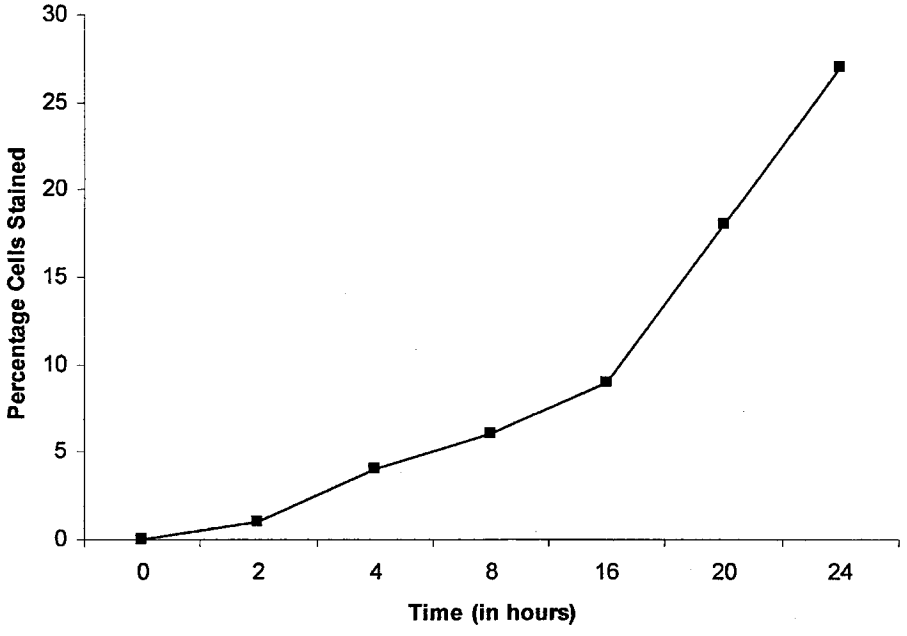


Figure 3.1. Expression of RSV protein over time in A549 cells infected with pRSV. A549 cells were plated and infected with pRSV at an MOI of 0.5. At the time shown they were fixed and stained with anti-RSV polyclonal antibody specific peroxidase.

3.2.2

IL8 activating cytokines detected in unpurified but not purified RSV preparations

I sought to test the hypothesis that the culture of RSV in Human HEp-2 laryngeal epithelial (HEp-2) cells induced the expression of cytokines, some of which had the potential to activate *IL8*. Therefore the purity of both the upRSV and pRSV preparations was tested with ELISA assays for known *IL8*-activating cytokines. Whilst 165 pg/ml of *IL8* was measured no IFN-γ, TNF or IL-1β were detected in the sucrose purified viral preparation. In contrast, upRSV samples grown on HEp-2 cells showed evidence of IL-8 (7403 pg/ml), TNF (16 pg/ml) and IL-1β (6 pg/ml).

3.2.3

IL-8 production significantly different between upRSV and pRSV preparations

Having established that cytokines with the potential to activate *IL8* were present in the upRSV samples but not in the pRSV preparation we sought to investigate whether this affected the IL-8 production from A549 cells. A549 cells were incubated with unpurified and purified live RSV as well as UV-inactivated unpurified and purified RSV after which the IL-8 concentration in the supernatant was measured at different time points. As shown in Table 3.1 over the time course unpurified RSV (upRSV) produced a fold increase of 63 in IL-8 production (554ng/ml) over cells incubated in media alone. In contrast, purified RSV (pRSV) induced only a 3.3 fold increase (27.4ng/ml) in IL-8 production. Therefore IL-8 production between purified and unpurified RSV preparations was significantly different. Moreover, UV-inactivated unpurified RSV also produced 19 fold more IL-8 (516ng/ml) than the purified RSV preparation. In contrast, inactivated purified RSV produced no more IL-8 (mean 6.68ng/ml) than media controls. These data support earlier work suggesting that

contaminating cytokines in the unpurified RSV preparation induce IL-8 production both independent of the virus and in synergy with it (Patel et al. 1998). Furthermore this phenomenon is not specific to IL-8 as Nitric Oxide is induced also by cytokines and RSV in a synergistic manner (Appendix 1).

Table 3.1 **IL-8 production after RSV infection**

	Mean IL-8 concentration (ng/ml)	*95% Confidence Interval*
upRSV	554	(223, 894)
upRSV UV-inactivated	516	(220, 812)
pRSV	27.4	(11.7, 43)
pRSV UV-inactivated	6.6	(3.1, 10.3)
Media alone	8.8	(5.8, 11.9)

IL-8 production at 72 hours from A549 cells after incubation with purified RSV (pRSV), unpurified RSV (upRSV), inactivated pRSV and inactivated upRSV at multiplicity of infection of 1. Observations are mean values from duplicate readings taken from four separate experiments.

3.2.4

pRSV infection of A549 cells produces IL8 activating cytokines

The induction of *IL8* activating cytokines from respiratory epithelial cells may prove to be a general phenomenon that is not restricted to HEp-2 cells. These cytokines may have a role in IL-8 production in A549 cells even after infection with pRSV. In order to determine the time course of *IL8*-activating expression from A549 cells infected with pRSV the concentrations of TNF, IL-1β and IFNγ were assayed in the supernatant of cells cultured in media alone, cells infected with purified RSV and from cells treated with inactivated RSV. TNFα, IL-1β and IFNγ were not detected in the supernatant in cells cultured in media alone or from cells stimulated with inactivated RSV. Whilst neither IL-1β nor IFNγ were detected from purified RSV treated A549 cell supernatants TNF was found (as shown in Figure 3.3.) at both 48 hours (mean 88.6 pg/ml, SE 51.6) and 72 hours (mean 165.9 pg/ml, SE 37.7).

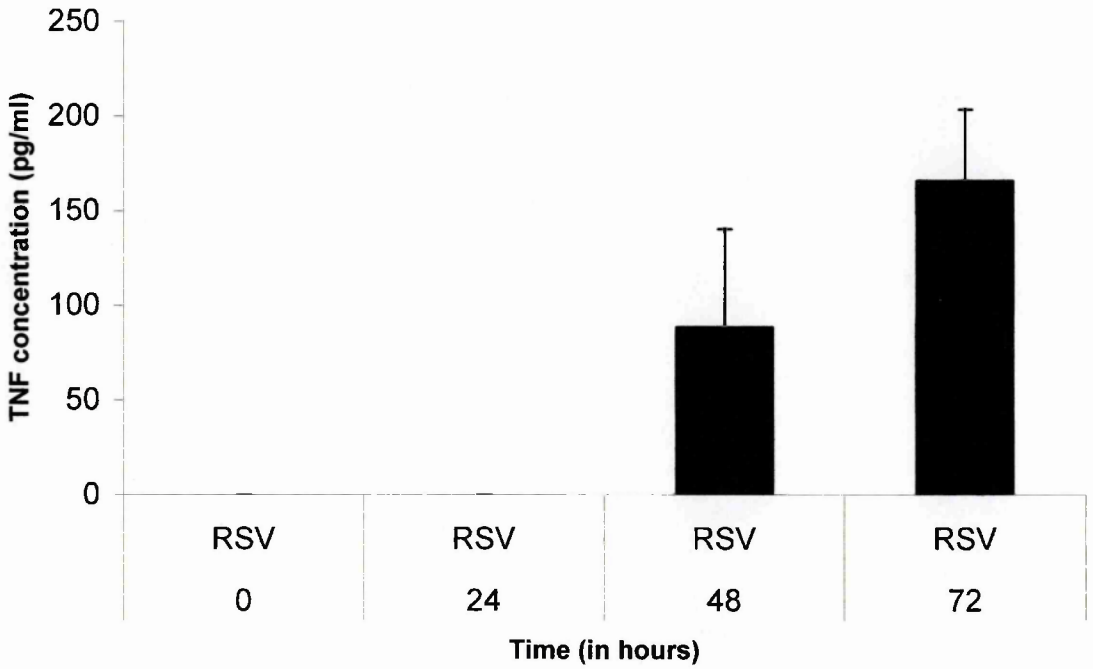


Figure 3.3. **Mean TNF concentrations over time in A549 cell supernatants after exposure to purified RSV.** Vertical bars represent the standard errors of the mean for duplicate readings from two separate experiments.

3.2.5

Comparison of IL8 transcription in response to TNF and pRSV stimulation

Having established that TNF is produced following pRSV infection in A549 cells we sought to confirm that TNF induced *IL8* and to compare the kinetics of *IL8* transcription between TNF and pRSV. A549 cells were exposed to 40 ng/ml of TNF or pRSV at an MOI of 1 and the RNA harvested at time points over 48 hours. As shown in the Northern blots of Figure 3.4 (a) and (b) *IL8* mRNA expression following TNF stimulation was first noted at half an hour after stimulation and was at a maximum within 2 hours. This *IL8* mRNA expression was sustained over the 48 hours of the experiment. By contrast, *IL8* mRNA expression following pRSV infection was first noted at 4 hours (Figure 3.4 (c) and (d)) and increased over time to reach a maximum level of transcription at the 36 and 48 hour time points.

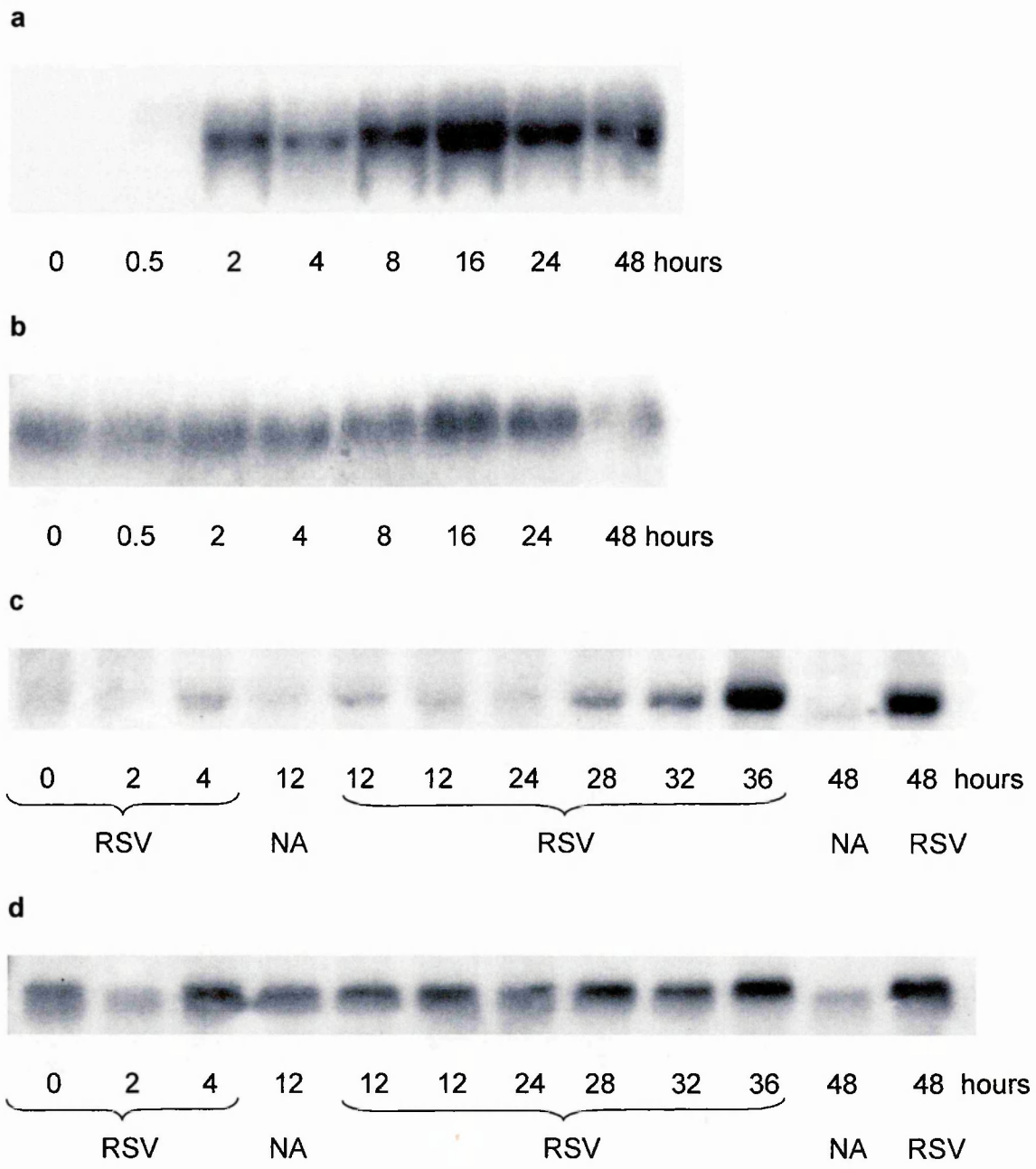


Figure 3.4. **Northern Blot analysis over time of *IL8* after TNF stimulation and pRSV stimulation.** *IL8* (**a**) and β-actin (**b**) mRNA expression from A549 cells treated with TNF at 40 ng/ml with *IL8* (**c**) and β-actin (**d**) mRNA expression from A549 cells either not activated (NA) or treated with pRSV at an MOI of 1. TNF induced *IL8* transcription performed by Dr Helen Fox and shown with permission from the author.

Discussion

3.3.1

Differences in IL-8 protein production in A549 cells following upRSV and pRSV stimulation

I have demonstrated that within the Kwiatkowski laboratory A549 cells are readily infected by pRSV. The kinetics of viral replication described here are consistent with previous reports as Garofalo and colleagues noted that the RSV F protein was present in the cytoplasm 8 hours after infection (Garofalo et al 1996). I have shown that the pRSV preparation is free from *IL8* activating cytokines where as the upRSV samples are not. I have demonstrated that upRSV infection of A549 cells produces a statistically significant increase in IL-8 production when compared with pRSV infection. The dose response relationship described here, and in previous reports (Fielder et al 1995), between either active or inactive upRSV and IL-8 implies that there are additional activators of *IL8* present in the upRSV samples other than RSV. Moreover, I have confirmed that pRSV infection of A549 cells produces other inflammatory mediators other than IL-8 such as Nitric Oxide (Appendix I) and TNF.

Laboratories using upRSV could not detect IL-1β or TNF in initial upRSV stock preparations grown on HEp-2 cells using an ELISA assay with detection limits of 20 pg/ml (Noah et al. 1993; Fiedler et al. 1995). However, subsequent studies have shown that pRSV infection of respiratory epithelial cells did produce a number of inflammatory mediators which included IL-1α (121 pg/ml), IL-1β (19 pg/ml) and TNF (77 pg/ml) (Patel et al. 1995). IL-1α, IL-1β and TNF are all known activators of *IL8* and it has been reported that

as little as 10 pg/ml of TNF or IL-1β could induce both *IL8* and *ICAM1* expression (Patel et al. 1995, Patel et al. 1998). Furthermore autocrine and synergistic up-regulation of *IL8* expression by IL-1α, IL-1β and TNF in pRSV-infected A549 cells has been described (Patel et al. 1998). Taken together these reports imply that cytokine contamination of upRSV samples could be functionally significant for *in vitro* studies of A549 cells.

3.3.2

IL8 transcription following upRSV, pRSV and TNF stimulation

Differences in the kinetics of cytokine production from A549 cells have been noted between unpurified and purified RSV. In common with TNF stimulation of A549 cells (Fiedler et al. 1998; Jamaluddin et al. 1998) upRSV has been reported to induce early *IL8* mRNA and protein expression at 2 hours (Fiedler et al. 1995) whereas purified RSV first induces *IL8* mRNA and protein from 6 hours (Garofalo et al. 1996). The pattern of *IL8* mRNA production described here is consistent with previous reports of pRSV-induced activation as transcription starts at low levels by 4 hours and gradually rises to peak by 36 hours.

NF-κB is an essential transcription factor for the transcriptional activation of *IL8* (Mastronarde et al. 1998; Casola et al. 2000). These distinct patterns in IL-8 protein expression in response to upRSV and pRSV are consistent with differences in the kinetics of NF-κB binding between the two stimuli. Activation with upRSV causes NF-κB activation at 15 minutes (Fiedler et al, 1996) a pattern which is similar to that described following TNF

stimulation (Fiedler et al. 1998; Jamaluddin et al. 1998). In contrast, pRSV produced a slower rate of RelA activation which peaked by 3 fold at 24 hours (Jamaluddin et al. 1998).

Differences in the electrophoretic gel mobility shift assay binding patterns were seen between TNF and pRSV stimulated A549 cells. When stimulated with recombinant TNF both C1 and C2 complexes were present. The C1 band predominated and consisted of RelA homodimers whereas C2 was supershifted by antibodies to both RelA (p65) and NF-κB1 (p50) (Jamaluddin et al. 1998). In contrast, pRSV also induced C1 and C2 complexes but in this instance C2 predominated (Garofalo et al. 1996).

A detailed study of the *IL8* gene promoter elements required for activation provided more evidence for differences between TNF and pRSV induction of *IL8* (Casola et al. 2000). TNF stimulated reporter genes when the preceding promoter contained only multimers of the NF-κB site. In contrast, pRSV induced activity is reconstituted by an RSV response element (RSVRE) (-162 to –132 nt), an AP-1 binding site (-132 to –99) and an NF-IL6 site in addition to the NF-κB sites. This implies that there are very different protein to protein as well as DNA to protein interactions required for RSV induced *IL8* activation as opposed to TNF activation.

Both RSV and TNF cause translocation of NF-κB to the nucleus. The ability of TNF to induce transcription could be explained by differences in REL family subunits (p65 homodimers predominating over p65-p50 heterodimers) or through post-translational modifications of these subunits. For instance, both TNF and IL-1 can stimulate phosphorylation of p65 (Zhong et al. 1997; Wang and Baldwin 1998) which is associated

with both an increase in association with co-activators (such as p300/CPB) as well as an increase in transcription (Zhong et al. 1997). Given that TNF and IL-1 are present in upRSV preparations, post-translational modification of REL family transcription factors may provide some of the explanation for the differences in *IL8* induction between upRSV and pRSV.

I recognize that *in vivo* RSV infection of epithelial cells induces the expression of a number of cytokines which may act in concert to enhance *IL8* expression. Even within an *in vitro* model I have shown that the pRSV induces late TNF production which may go on to augment *IL8* expression (Patel et al. 1998). However, the complexity of the interactions present between RSV and the *IL8* activating cytokines present in the upRSV preparation would be great. The upRSV preparation may prove to be too heterogeneous a stimulus to use when attempting to dissect how RSV affects the functional genetics of SNPs within the *IL8* gene.

3.3.3

Concluding Points

- Purified RSV made within this laboratory infected A549 respiratory epithelial cells resulting in RSV protein production and syncytia formation.
- TNF and IL-1β, which are known to be *IL8* activating cytokines, were detected in unpurified but not purified RSV preparations.
- Activation of *IL8* is considerably greater when unpurified RSV is used as compared to purified RSV.
- TNF is produced on purified RSV infection of A549 cells.
- TNF induces sustained and profound *IL8* transcription.
- Synergistic activation of *IL8* by the combined action of RSV and *IL8* inducing cytokines is likely in RSV-induced bronchiolitis.

CHAPTER FOUR

IN VIVO HAPLOTYPE-SPECIFIC EXPRESSION OF *IL8* IN PRIMARY CELLS

4.1.1

Introduction

Traditionally, allele-specific variation in gene expression has been associated with epigenetic phenomena such as X-chromosome inactivation and genomic imprinting. However, it has become apparent that gene expression in non-imprinted autosomal genes varies greatly between individuals (Cheung et al. 2003), populations (Oleksiak et al. 2002) and species (Enard et al. 2002). Moreover, through experiments conducted in mice it appears that allele-specific expression patterns are inherited (Cowles et al. 2002). These findings have served to underlie the potential importance of natural variation in non-coding DNA sequences in governing gene regulation and subsequent phenotypic diversity (Knight 2004).

There is evidence that DNA sequence variation in the form of SNPs does result in allele-specific differences in expression in mice (Cowles et al 2002). Cowles and co-workers studied F1 crosses of four in-bred mice and showed on a screen of genes expressed in liver, spleen and brain that 7 out of 69 genes showed allele-specific expression. These results were verified in the four mouse strains by examining other SNPs within the transcript of the 7 genes. Allele-specific expression was confirmed in 4 of the genes where differences in expression above 1.5-fold were observed but not in the 3 which had lower degrees of

differential transcription. Of considerable interest is the observation that allele-specific expression is frequently tissue specific in nature. Cowles and colleagues noted that two of the four genes under study were allele-specifically expressed in the liver but not in the brain or the spleen (Cowles et al. 2002).

However, the number of genes which are estimated to have allele-specific expression is likely to be influenced by both the sensitivity of the detection technique and the species in which the study is conducted. For instance, Cowles et al (2002) used a fluorescently labeled dideoxynucleotide primer extension technique which could measure a 1.5-fold difference or greater in expression between alleles. Studies of humans have used three different techniques namely, the Affymetrix HuSNP chip system (Lo et al. 2003), a fluorescent dideoxy terminator-based method (Yan et al. 2002) and a primer extension with Genescan detection technique (Bray et al. 2003a). For each method the threshold sensitivities for differences in fold expression were 2.0, 1.2 and 1.2 respectively. All three techniques estimate that the proportion of allele-specifically expressed genes in humans is higher than that seen in mice. A study of 602 genes using an Affymetrix HuSNP chip system concluded that in at least one of the 7 individuals allele-specific expression occurs in 54% of genes. It remains to be seen whether the Affymetrix HuSNP chip system is a sufficiently accurate measure of allele-specific expression (Knight 2004). Analysis of human lymphocytes and brain tissue using the fluorescent dideoxy terminator-based method (Yan et al. 2002) showed that a more modest 46% of genes were allele-specifically expressed. The same proportion of allele-specifically expressed genes was seen with a study using the primer extension with Genescan detection technique on brain tissue (Bray et al. 2003a).

Taken together these data suggest that mammalian allele-specific expression of genes is likely to be both common and context specific. Given the genetic epidemiological evidence suggesting that *IL8* had a role in RSV-induced bronchiolitis disease severity I sought to investigate whether *IL8* showed allele-specific expression in primary respiratory epithelial cells. As contextual factors appear to be important to allele-specific transcription I sought also to investigate whether any such differential expression was tissue specific by looking at IL-8 production in primary lymphocytes. However, at present there is little understanding of what level of allele-specific expression is likely to be physiologically important. I reasoned that relatively small differences in the expression of a chemokine with potent properties of neutrophil chemotaxis could have an effect on disease outcome. In order to detect these small differences I sought to measure *IL8* allele-specific expression using two different techniques namely, mutant analysis by PCR and restriction enzyme cleavage (MAPREC) and primer extension mass spectrometry (PE/MS).

Results

4.2.1

Optimisation of MAPREC as a method for haplotype-specific quantitation

PCR primer [1] was designed so as to introduce a *G* residue, instead of the normal *A* residue, into the PCR product. For example in the +2767T allele forward strand the primer would change the normal sequence (the SNP being marked in bold) from 5'-TCAG*A*TA**T**AC-3' to 5'-TCAG*G*TA**T**AC-3'. This had the effect of introducing a *Bst* Z 171 restriction site (5'-GTATAC-3'; where the cut occurs at the underlined A residue) into the reverse strand in the presence of the +2767T allele (+2769 5'-GTATACCTGA-3' +2760) but not the +2767A allele (+2769 5'-GTTTACCTGA-3' +2760). Therefore restriction digestion of the amplified PCR product would produce an uncut 129 base pair (bp) fragment in the case of the +2767A allele but 103 bp and 29 bp fragments where the +2767T allele is present.

To ensure that the +2767T and +2767A template alleles had equal amplification the PCR reaction was terminated at a cycle number where the amplification process was still in the log phase (Figure 4.1). The optimal cycle number was then set between 28 to 30 cycles.

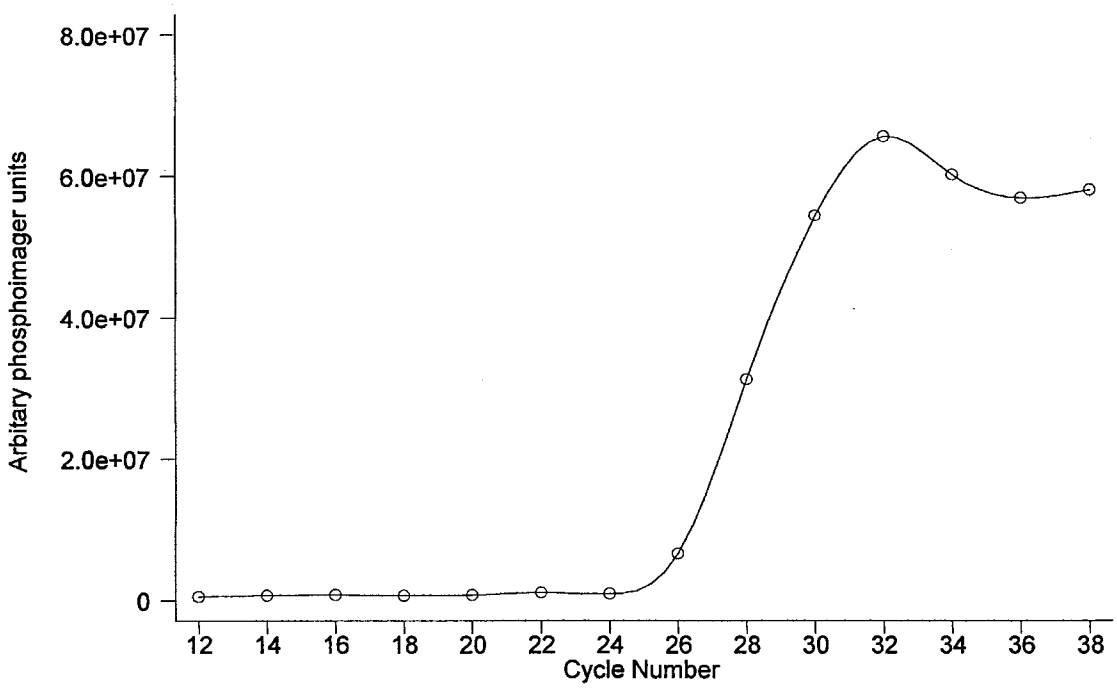


Figure 4.1

Quantification of product amplification to PCR cycle number. The PCR reaction used genomic DNA and [α-^{32}P]dCTP labeled nucleotides with samples taken every other cycle. The PCR samples were electrophoresed on a non-denaturing polyacrylamide gel and the amplification product quantified using the Phosphoimager. The graph shows the log phase for product amplification occurred between 26 and 30 cycles.

The restriction digest would be unable to cut the PCR products amplified from the +2767A allele but would also be unable to digest heterodimers (for example +2767T Forward (5'-TCAGGTA**T**AC-3') with +2767A Reverse (5'-GT**T**TACCTGA-3'); and +2767A Forward (5'-TCAGGTA**A**AC-3') with +2767T Reverse: (5'-GT**A**TACCTGA-3'); the SNP is marked in bold) formed through incorrect annealing from the previous PCR cycle. This has the potential to bias the quantitation of the PCR products towards the +2767A allele. To ensure that does not occur the [α-^{32}P]dCTP was added in the last cycle only so that there would be no further denaturation and annealing after the incorporation of the radioactive signal to the PCR products.

To test the sensitivity of the PCR and the quantification of the radiolabeled digested products the technique was tested on a range of standards. Genomic DNA homozygous for either the +2767A or the +2767T allele was mixed with concentrations of the other allele at ratios (for +2767A/+2767T) of 0.5, 0.8, 1.0, 1.2 and 2.0 (Figure 4.2). After PCR and restriction digest over three independent and consecutive experiments the following mean values were obtained: 0.58 (95% CI 0.52, 0.64); 0.88 (95% CI 0.79, 0.97); 0.99 (95% CI 0.93, 1.07); 1.18 (95% CI 1.1, 1.27) and 1.7 (95% CI 1.59, 1.81). Therefore, the PCR and restriction digest were sensitive enough to detect a significant difference between DNA samples of 20% or above.

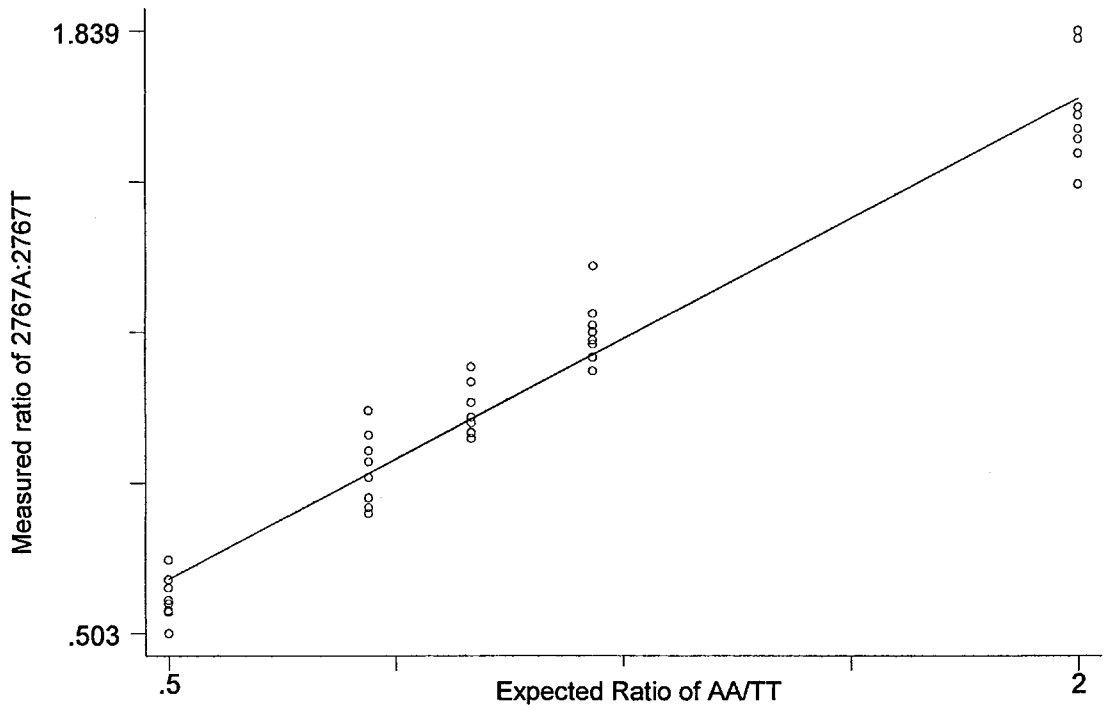


Figure 4.2

Standard Curve for haplotype-specific radiolabeled PCR. Standard curves were produced for the +2767 SNP by preparing different ratios of genomic DNA (shown on the x axis) from two individuals, homozygous AA or homozygous TT at that polymorphism. The observed ratios between the A and the T alleles were then measured by radiolabelled PCR and restriction digest (shown on the y axis). These data were derived from duplicate readings for three independent experiments using concentration ratios of homozygous genomic DNA (+2767AA over +2767TT) at 0.5, 0.8, 1.0, 1.2 and 2.0. After radiolabelled PCR and restriction digest the following mean values were observed: 0.58, 0.88, 0.99, 1.18 and 1.7. The regression co-efficient is 0.71 (standard error 0.026), R-squared = 95%, $p<0.001$.

4.2.2

Harvest, culture and stimulation of primary respiratory epithelial cells

30 healthy adult volunteers were recruited from day case surgery operation lists and agreed to have primary respiratory epithelial cells harvested from their nasopharynx under general anaesthetic. Consent to approach patients for anonymous tissue collection was obtained from the Oxfordshire research ethics committee (OXREC) number 99/5/27. The primary cells were incubated in media containing TNF (5ng/ml) and harvested at 3 or 6 hours. Venous blood was taken for genotyping for both the -251 and +2767 SNPs and as a source of genomic DNA. Of the 30 patients consented 8 individuals (2 men and 6 women) were heterozygote for both the -251 and +2767 SNPs. From these 8 patients the mRNA was extracted and the complementary DNA (cDNA) synthesized with no detectable contamination by genomic DNA. Maximal IL-8 production from primary epithelial cells had previously been observed by 3 hours after stimulation with 30ng/ml of TNF (J Hull personal communication). Stimulation with purified RSV resulted in lower level IL-8 production which, given the relatively small number of primary cells, was not sufficient for haplotype-specific expression analysis.

4.2.3

Haplotype-specific expression of IL8 not detected by MAPREC

The +2767 polymorphism is present in the mature mRNA. The quantification of the ratio of the +2767A over the +2767T cDNA was obtained by PCR radiolabelled amplification of the cDNA and restriction digest. As a control the ratio of the cDNA was compared to that of the genomic DNA as the latter should have a ratio of close to 1. As shown in Figure 4.3 there was not a significant difference seen between the +2767A/+2767T ratio for cDNA

compared to genomic DNA for cells incubated for 3 hours. However, the trend was for more of the +2767T haplotype to be transcribed than the +2767A haplotype. When the +2767A/+2767T ratio for cDNA and genomic DNA was compared in cells incubated for 6 hours (Figure 4.4) again no significant differences were seen between the two ratios. In this case the trend was in the opposite direction with greater transcription of the +2767A haplotype.

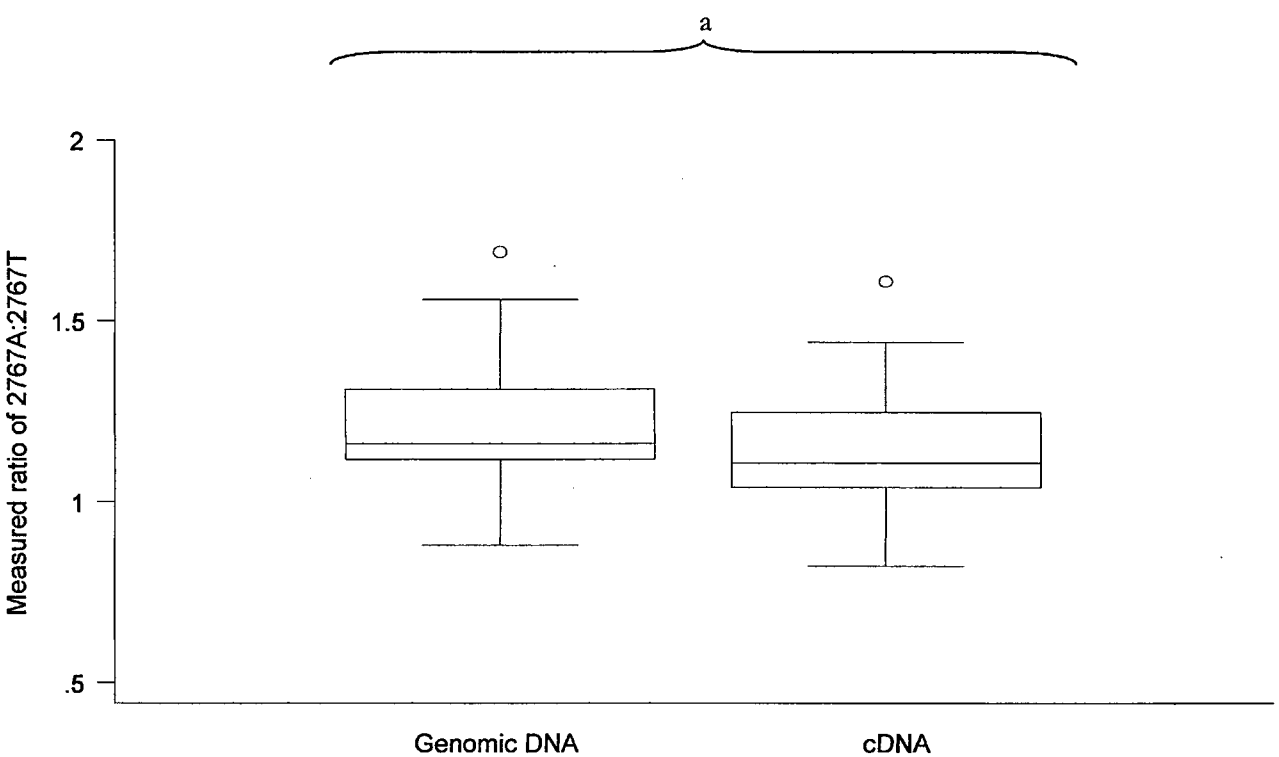


Figure 4.3

Allele specific PCR of *IL8* Genomic and cDNA from primary respiratory epithelial cells after TNF stimulation for 3 hours. The 2767A/2767T mean ratio for genomic DNA was 1.23 (95% CI 1.12, 1.33) compared to 1.15 (95% CI 1.05, 1.25) for cDNA. All data are reported as the median increase (with inter-quartile range (IQR)) in eight individuals from two separate experiments. [a]Paired *t* test $p = 0.06$.

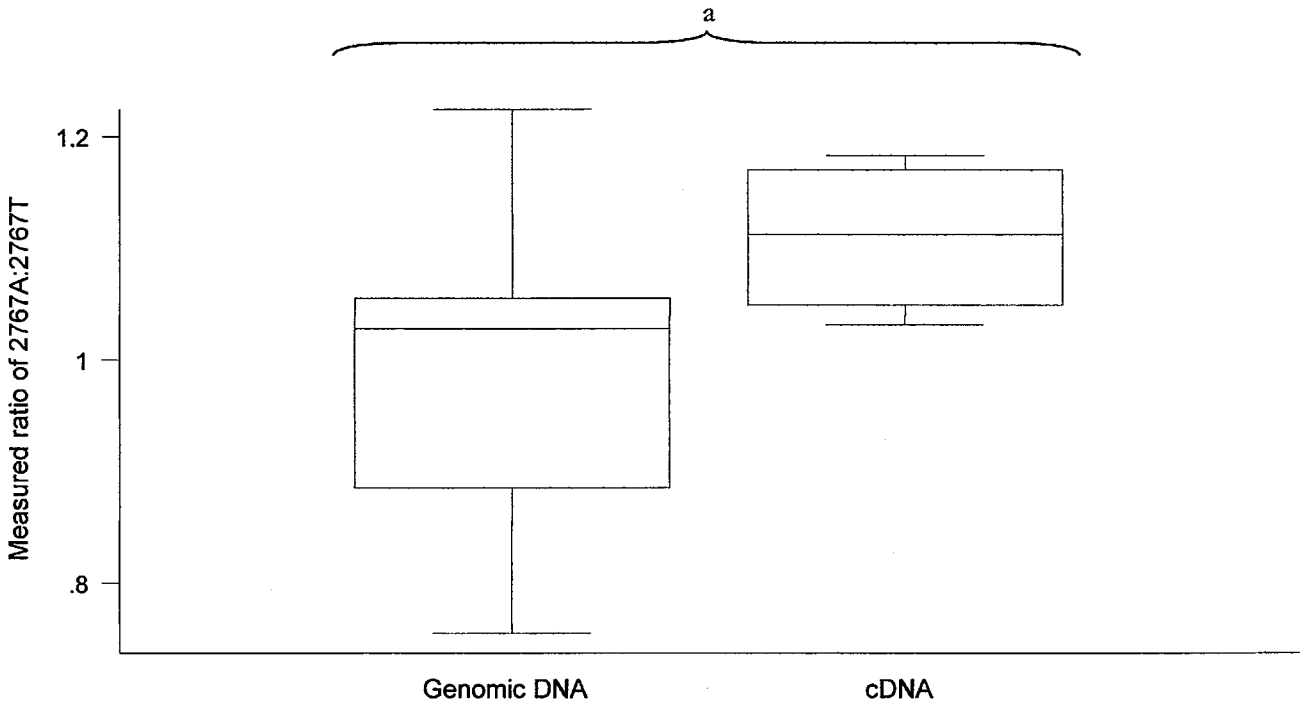


Figure 4.4

Allele specific PCR of *IL8* Genomic and cDNA from primary respiratory epithelial cells after TNF stimulation for 6 hours. The 2767A/2767T mean ratio for genomic DNA was 0.99 (95% CI 0.87, 1.11) compared to 1.11 (95% CI 1.06, 1.17) for cDNA. All data are reported as the median increase (with inter-quartile range (IQR)) in eight individuals from two separate experiments. [a]Paired *t* test p = 0.09.

4.2.4

Optimisation of PE/MS as a method for haplotype-specific quantitation

As before the +2767 polymorphism was used to assess the relative abundance, as measured by PE/MS (Knight et al 2003), of the transcripts derived from each of the two haplotypes. To assess the accuracy of the PE/MS method with the +2767 SNP we prepared a standard curve by mixing genomic DNA at different ratios from two individuals who were homozygous for the +2767A and +2767T alleles (Figure 4.5). We observed a close linear relationship between the ratio of the alleles and the observed ratio of peak areas (r^2 0.95, p=0.0000). The SNP +2767 shows a systematic difference in signal intensity from the expected value of 1:1 with genomic DNA. Since we expect the genomic DNA to represent a 1:1 ratio between alleles the average difference from 1 in the genomic samples was used to provide a correction factor that was applied to both the genomic and cDNA samples for each experiment as previously demonstrated (Bray et al 2003a).

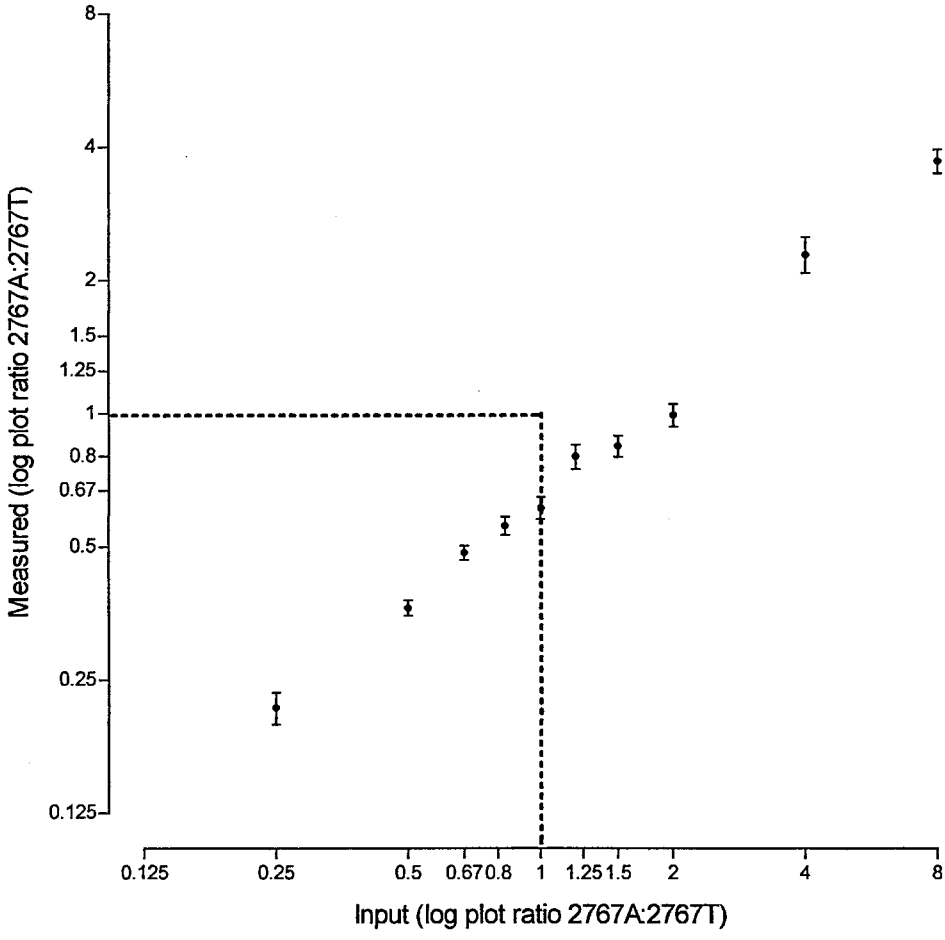


Figure 4.5

Standard Curve for haplotype-specific radiolabeled PCR. Standard curves were produced for the +2767 SNP by preparing different ratios of genomic DNA (shown on the x axis) from two individuals, homozygous AA or homozygous TT at that polymorphism. The observed ratios between the A and the T allelic peaks were then measured by PE/MS (shown on the y axis).

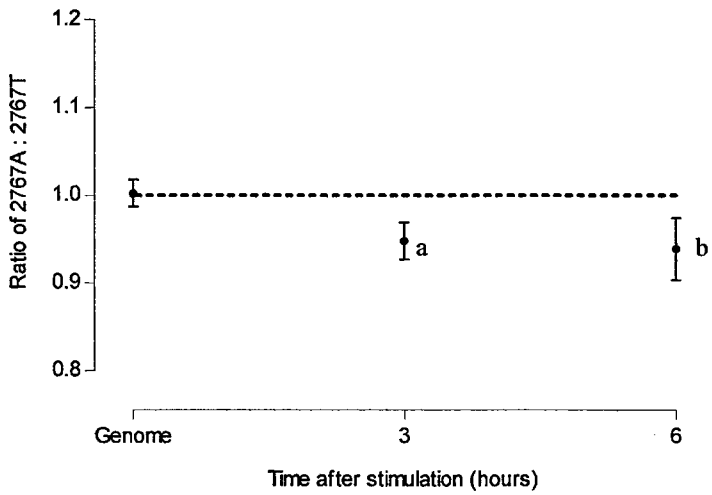
4.2.5

Haplotype-specific expression of IL8 detected by PE/MS

The abundance of the two alleles from the +2767 SNP transcribed marker was quantified by PE/MS. For each individual, the ratio of the +2767A allele (haplotype 1) over the +2767T allele (haplotype 2) in the cDNA and genomic DNA was compared. The relative abundance of the haplotype 2 transcript when compared with that of haplotype 1 transcript was significantly increased over time reaching a ratio of 0.94 at 6 hours (p=0.037) (Figure 4.6a). To explore whether the *in vivo* allele-specific difference demonstrated in primary respiratory epithelial cells was present in other primary cells, we harvested primary lymphocytes from six individuals heterozygote for haplotypes 1 and 2. After stimulation with LPS we observed no significant increased transcription for one *IL8* haplotype over another (Figure 4.6b).

We anticipated that the mechanism for the haplotype-specific expression demonstrated here might require exploration through *in vitro* studies of plasmid reporter genes and DNA-protein interaction. As these *in vitro* studies would take place within immortalized cells lines we sought to determine what patterns of *IL8* haplotype-specific expression occurred in this cell type. In contrast to our findings with primary cells, human H460 pulmonary epithelial cells stimulated with TNF showed a significant increase in production of the Haplotype 1 (Figure 4.7) in both non-activated cells (at 0 hours ratio 1.075, p=0.009) and activated cells (at 6 hours ratio 1.11, p=0.0001).

a



b

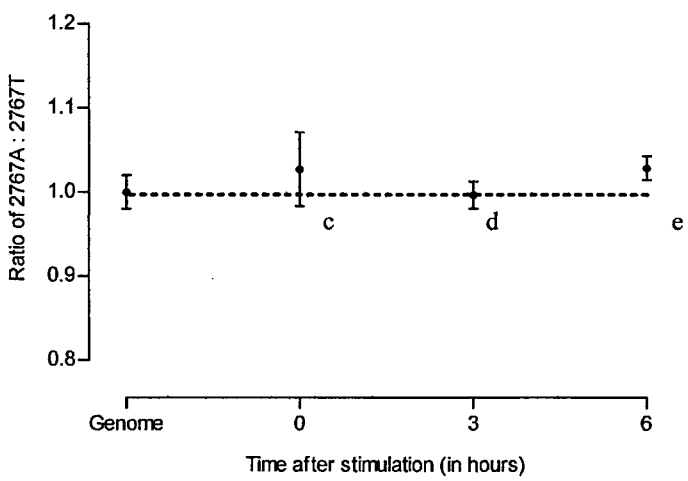


Figure 4.6 **Haplotype-specific transcription of *IL8* in primary epithelial cells and in primary lymphocytes.**

Figure 4.6 **Haplotype-specific transcription of *IL8* in primary epithelial cells and in primary lymphocytes.** After incubation in media containing the appropriate stimulant (**a**, 30ng/ml TNF for Primary Nasal Epithelial cells; **b**, 10 μg/ml LPS for Primary Lymphocyte cells) mRNA was extracted at the times shown and analyzed by RT-PCR using primers specific for the +2767A and +2767T polymorphisms. For each individual, cDNA and genomic DNA were amplified in four independent PCR reactions which in turn were each spotted as four replicates on the spectroCHIP before PE/MS. After correction for systematic difference in signal intensity the +2767 allelic ratio the genomic DNA was compared with the cDNA for each individual by a paired *t*-test. Genome represents the +2767 allelic ratio for genomic DNA. Values are expressed as the mean ratio ± the standard error of the mean of the +2767A allele over the +2767T allele. (**a**) At 3 hours 0.95, ^{a}p=0.067; at 6 hours 0.94, ^{b}p=0.037. (**b**) At 0 hours 1.03, ^{c}p=0.67; at 3 hours 1.0, ^{d}p=0.34; at 6 hours 1.03, ^{e}p=0.31.

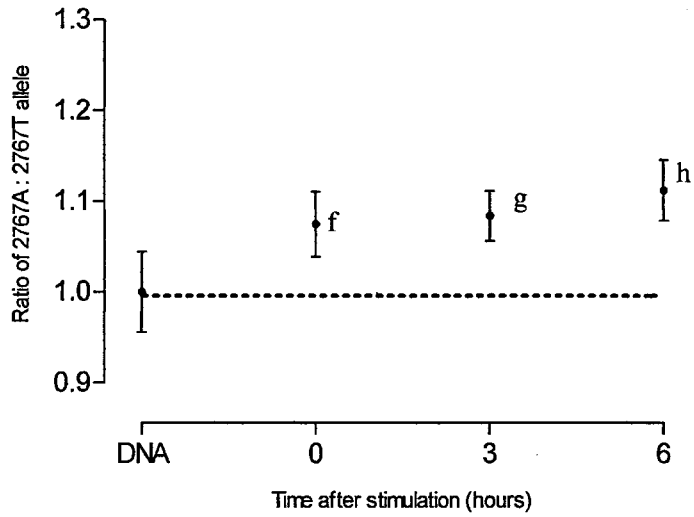


Figure 4.7 **Haplotype-specific transcription of *IL8* in Human H460 pulmonary epithelial cells.** After incubation in media containing the appropriate stimulant (30ng/ml TNF for Human H460 pulmonary epithelial cells) mRNA was extracted at the times shown and analyzed by RT-PCR using primers specific for the +2767A and +2767T polymorphisms. For each individual, cDNA and genomic DNA were amplified in four independent PCR reactions which in turn were each spotted as four replicates on the spectroCHIP before PE/MS. After correction for systematic difference in signal intensity the +2767 allelic ratio the genomic DNA was compared with the cDNA for each individual by a paired *t*-test. Genome represents the +2767 allelic ratio for genomic DNA. Values are expressed as the mean ratio ± the standard error of the mean of the +2767A allele over the +2767T allele. At 0 hours 1.075, ^{f}p=0.009; at 3 hours 1.08, ^{g}p=0.002; at 6 hours 1.11, ^{h}p=0.0001.

4.2.6

Karyotype and Fluorescent in situ hybridization (FISH) for IL8 in Human H460 pulmonary epithelial cells

In order to explain the marked differences in the haplotype-specific expression of *IL8* between primary cells and the H460 pulmonary epithelial immortalized cell line we enlisted the assistance of Emmanula Volpe and Andrew Jefferson of the Molecular Cytogenetics Department of the Wellcome Trust Centre for Human Genetics. As expected from the American Tissue Culture Collection (ATCC) description the H460 cell line had a hypo-triploid karyotype. When analysed by FISH for chromosome 4 number, this cell line shows two apparently normal number 4 chromosomes in 69% of cells whilst in 12% of cells there was either only one or more than two number 4 chromosomes. In the remaining 19% of cells the result was not clear. FISH analysis for the *IL8* gene copy number shows that of 33 metaphases analysed, all of them presenting two chromosome 4 homologues, 30% had two copies of *IL8* whilst in 45% of cells there was just one copy. In 24% of cells the *IL8* copy number was uncertain. We could not easily determine *IL8* haplotype in the cells where there was only copy. In conclusion, this FISh study appears to show that a large proportion of cells in the H460 cell line were not heterozygote for the *IL8* haplotypes of interest. This finding means that the results of any *IL8* haplotype-specific experiments in this cell line are impossible to accurately interpret.

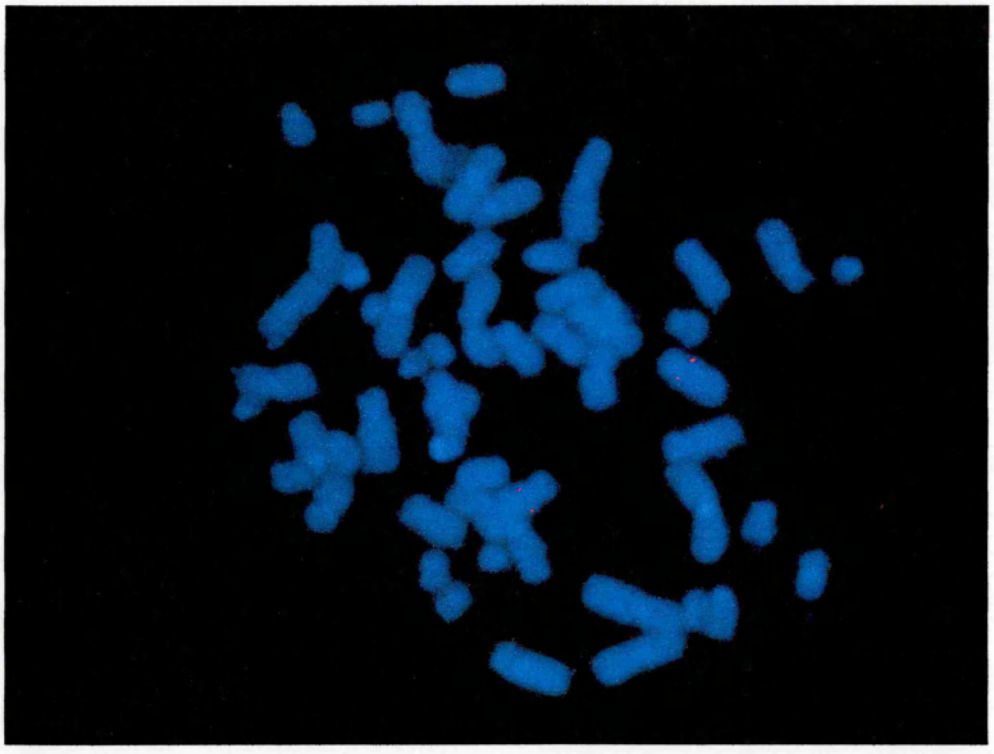


Figure 4.8

Karyotype and FISH probe for the *IL8* gene

Representative photomicrograph of H460 chromosomes counterstained with DAPI (blue) to which a FISH IL8 cDNA probe (red) has been added. Karyotype and FISH probe conducted by Andrew Jefferson and Emmanula Volpi and is reproduced here with their permission.

Discussion

4.3.1

The biological significance of the IL8 haplotype-specific transcription by PE/MS

I sought to investigate the transcriptional regulation of *IL8* through examining the haplotype-specific differences in mRNA expression between the two main haplotypes described in the Caucasian population. Previously it has been shown that the -251A SNP was associated with susceptibility to severe RSV-induced bronchiolitis (Hull et al. 2000) and that this association was shown to include the SNPs comprising haplotype 2 (Hull et al 2001) (Figure 1.6). Using PE/MS alone I have shown that in primary respiratory epithelial cells heterozygous for haplotypes 1 and 2, mRNA derived from haplotype 2 was detected significantly more than that from haplotype 1. This might occur through either increased transcription or through increased mRNA stability of haplotype 2 as compared with haplotype 1. This is one of the few studies demonstrating that a gene associated with disease susceptibility is subject to *in vivo* allele-specific expression in primary cells (Bray et al. 2003b). The increased abundance of mRNA from haplotype 2 is compatible with the observation that the concentration of *IL8* is correlated with severe disease (Smyth et al 2002) and that severely affected individuals have increased neutrophil infiltration in the lungs (Everard et al. 1994, Aherne et al 1970).

Though the size of the difference in mRNA detected between the two haplotypes was modest (6%) it remains statistically significant and may be of biological importance for the following reasons. IL-8 is a potent chemotactic agent for neutrophils in humans and the haplotypes we have studied occur frequently (Figure 1.6). Small differences in the levels of gene expression could have important effects on disease outcome and affect a high

proportion of the population. The *IL8* mRNA production measured here occurred in the context of primary respiratory cells isolated from other cell types such as neutrophils and macrophages. Both neutrophils and macrophages are known to play an important role in the host response to RSV infection and may influence the degree of haplotype-specific expression of *IL8* that occurs in epithelial cells.

Moreover, the effect of the haplotype on mRNA abundance appeared to be cell specific as it was observed in primary respiratory epithelial cells but was not seen in primary lymphocyte cells. However, the primary respiratory epithelial cells and the primary lymphocytes were derived from different populations. The proof for the localization of haplotype-specific *IL8* transcription to one cell type could only come from experiments on a population of primary cells derived from different tissues in the same individual. In addition the degree of allele-specific *IL8* expression may also be limited by developmental stage and environmental factors. The presence of marked allele-specific expression of *IL8* in infants but not adults could be an additional factor to explain the severity of RSV-induced bronchiolitis in this age group.

IL8 lies within a chemokine cluster that includes CXCL1, CXCL2, CXCL3, CXCL5, CXCL6, pro-platelet basic protein (PPBP) and Platelet factor-4 (PF4) (Modi and Chen, 1998). Moreover, PF4, PPBP and IL-8 all have chemotactic properties in relation to lymphocytes (Wuyts et al 1994). It is possible that *IL8* is only one of a number of genes in that chemokine cluster that exhibit haplotype-specific expression and it is the combined effect of these chemokines which modifies RSV-induced bronchiolitis disease-susceptibility. The Kwiatkowski laboratory has addressed this question recently by attempting to localize

the bronchiolitis disease association to one or more of the genes in that chromosomal location. (Hull et al. 2004). Three major clusters of high linkage disequilibrium were identified that contained a novel gene RASSF6, *IL8* and the remaining genes of the chemokine cluster respectively. Only markers within the RASSF6 and *IL8* haplotype blocks appeared to show a disease association to RSV-induced bronchiolitis. Further analysis of the data shows that the effect depends on the presence of both elements. The RASSF6 gene is of considerable interest as it contains a Ras association domain. Intriguingly, RSV infection has been shown to activate RhoA, a member of the Ras superfamily, which in activated state can induce *IL1β*, *IL6* and *IL8* (Gower et al. 2001). Moreover, preliminary experiments that I have conducted appear to show allele-specific expression of RASSF6 in lymphocytes. Technical difficulties and time constraints prevented the optimization of these allele-specific experiments in primary respiratory cells.

I acknowledge that we have assumed that the primary mechanism by which IL-8 affects susceptibility to RSV-induced bronchiolitis is through transcription. Whilst studies suggest that differences in gene expression levels account for a major part of the variation within and between species (Johnson and Porter 2000, Levine 2002) it is possible that *IL8* does not fit this pattern. In addition to the modest allele-specific differences in *IL8* transcription shown here there could be more significant mechanisms of translational regulation which may contribute to RSV-induced bronchiolitis disease susceptibility.

4.3.2

Methodological merits and limitations to the demonstration of haplotype-specific IL8 expression using cultured primary cells and PE/MS

The study design used cells heterozygous for the allele of interest allowing the detection of small variations in allele-specific transcription independent of other host, environmental and technical factors (Knight et al 2003, Yan et al 2002). For instance, the investigative approach would control for differences in the cell number and strength of the stimulus between experiments. The experimental design would also control for differences in the IL-8 receptor signaling and signal transduction pathways between individuals as both haplotypes of the *IL8* gene should be stimulated to the same degree. The host and environmental factors may be one reason why, like a number of other studies, I observed differences in haplotype-specific expression of *IL8* between individuals (Bray et al 2003a, Lo et al 2003, Yan et al 2002, Cowles et al 2002).

However, there are a number of technical limitations to the study design that require further consideration. Firstly, verifying the observed allele-specific expression of *IL8* by testing additional marker SNPs is limited as the +2767 SNP is the only transcribed marker. Application of other methods of *in vivo* allele specific expression, such as haploCHIP (Knight et al 2003), to certain types of primary cells may be impractical as it requires large numbers of cells to obtain sufficient chromatin for immunoprecipitation. Secondly, the systematic allelic difference in PE/MS signal intensity seen with the +2767 SNP using genomic DNA is well described for other polymorphisms (Knight et al 2003) and usually relates to biases in primer design (Bray et al 2001). Thirdly, I chose to use TNF as a stimulus for *IL8* activation and not RSV as purified RSV produces a smaller *IL8* response from

respiratory cells. This would have produced an insufficient quantity of *IL8* transcript for the allele-specific assay as modest numbers of primary respiratory cells were obtained from each patient at the time of harvest.

4.3.2

Methodological limitations to the demonstration of haplotype-specific IL8 expression using immortalized cell lines and PE/MS

There was a significant difference in the *in vivo* allele specific expression of *IL8* between primary respiratory cells and the immortalized H460 cell line. This may reflect differences in the gene regulation between the two cell types as *IL8* increases tumor growth through promoting angiogenesis in bronchogenic (Smith et al. 1994) and non-small cell lung carcinoma (Arenberg et al. 1996) and is constitutively expressed in these cancer cells (Arenberg et al. 1996). These factors highlight the potential limitations to studying allele specific expression of angiogenic chemokines in immortalized cell lines. In common with other cell lines derived from cancerous tissue the H460 cells have a hypo-triploid karyotype. In addition a significant number of the cells which contain two copies of chromosome four have only one copy of the *IL8* gene. Given the large structural rearrangements in the genome of these cells they proved to be a poor model system for studying haplotype-specific expression in heterozygote cells.

4.3.3

Comparison of MAPREC and PE/MS as a method for demonstrating haplotype-specific transcription

I have demonstrated a modest but significant degree of haplotype-specific transcription of *IL8* using PE/MS which was not found using MAPREC. The two techniques have been compared before in the analysis of mutations in live poliovirus vaccines (Amexis et al. 2001). In common with our findings Amexis and colleagues (2001) concluded that both techniques had excellent correlation, however MAPREC was seen to have a number of limitations.

For quantification of the products from each allele MAPREC requires the creation of an artificial digest site, radiolabel of the PCR products followed by digest and finally separation of the labeled DNA on a polyacrylamide gel. The use of hazardous materials in this procedure and the need for a digest step means that the experiment is both time consuming and difficult to scale up to incorporate multiple samples. For instance, within the Kwiatkowski laboratory it was safe and practical for a sole operator to manage up to a maximum of 50 samples in any one MAPREC experiment. By comparison PE/MS allowed for the PCR of each sample in four separate reactions which in turn after primer extension were spotted four times on the MALDI-TOF plate. Each of these spots was then laser ionized a further four times the mean of which was given as the peak height for each allele at that spot. Through this study design it was possible to minimize random variation from the allele quantification using PE/MS. In our laboratory a sole operator could comfortably handle up to 384 PCR reactions in an experiment over one day. Quality control of the PE/MS procedure has shown that the greatest variation occurs at the PCR amplification stage (Julian Knight and Kirk Rockett, personal communication).

Further analysis of the MAPREC technique has shown that the digest of the PCR products is 96 to 99% efficient in normal circumstances which may be insufficiently sensitive to detect differences in allele specific expression as small as 6%. However, the digest efficiency can be more severely affected where there is incorrect incorporation of nucleotides in the course of amplification by PCR as these incorrect bases may alter the restriction digest site. This has the potential to introduce significant bias to the allele-specific quantification. Incorrect base incorporation could also occur with PE/MS but would be easier to detect when using MALTI-TOF as the product containing the incorrect bases would have a different peak molecular mass to the one predicted for that allele.

Whilst MALDI-TOF is a fast and accurate method of allele specific quantitation (Tsuchihashi and Dracopoli 2002) as I have demonstrated here it can produce unequal signal intensities for alleles of the same SNP (Knight et al. 2003). These unequal signal intensities have been attributed to either preferential incorporation of one dideoxynucleotide over another at the primer extension step or to sequence specific differences in ionisation (Bray et al 2001). Both of these issues can be corrected or ameliorated by modification of the primer sequence or through changing the primer sequence to the complementary strand.

4.3.4

The possible mechanism for the IL8 haplotype-specific transcription

Identifying and understanding *cis*-acting regulatory units in humans presents a significant challenge. It is not possible from knowledge of sequence variation to predict which SNPs affect regulation. This difficulty in identifying functional SNPs is made worse by our limited ability to define the regulatory regions of most genes as these regions may be several kilobases from the transcriptional start site. Moreover, looking at variation in transcript levels for a gene from different individuals may not provide the answer as any differences in transcription may be due to *trans*-acting or environmental factors. Identification of regulatory regions that cause allele-specific expression can only be done by studying the two alleles under identical conditions. The study of gene transcription of *IL8* in heterozygous primary cells fulfills some of these conditions. However, these experiments do not allow us to define the extent of the regulatory regions governing *IL8* transcription nor do they help us identify whether the SNPs in the haplotype are merely markers or if they have a functional role. In the presence of a population structure that was not dominated by two mirror haplotypes it may have been possible to select SNPs on a number of different haplotypes that were associated with greater degrees of haplotype-specific expression *in vivo*. This question is addressed in Chapter 7 by using the *TNF/LTA* locus as a model system. In the absence of a population structure that suggests which SNPs are likely to be functional we have adopted *in vitro* techniques of plasmid transfection experiments and EMSA studies of DNA to protein interactions in Chapters 5 and 6 in an attempt to identify some aspects of the regulatory mechanism of *IL8* haplotype-specific expression.

4.3.5

Concluding Points

- Haplotype-specific expression of *IL8* was not detected by MAPREC in primary respiratory epithelial cells.
- Modest increased expression of the *IL8* disease severity associated haplotype-2 was detected by PE/MS in primary respiratory epithelial cells.
- Haplotype-specific expression of *IL8* not detected by PE/MS in primary lymphocyte cells.
- Karyotype and Fluorescent *in situ* hybridization (FISH) studies for *IL8* in Human H460 pulmonary epithelial cells shows that this cell line is a poor model for haplotype-specific expression in primary cells.
- The mirror haplotype structure of the *IL8* gene in a caucasian population prevents resolution *in vivo* of which SNPs in the haplotype show functional properties.

CHAPTER FIVE

INVESTIGATION OF THE 5' FLANKING REGION OF THE *IL8* GENE IN THE REGULATION OF *IL8* TRANSCRIPTION

5.1.1

Introduction

Since it was established that *IL8* is readily induced by a number of agents, including TNF, IL-1 and RSV, the 5' flanking region of the gene has been studied extensively so that *cis*-elements involved in activation and silencing could be better understood.

The first transfection experiments examining the 5' flanking region of the *IL8* gene used upRSV as a stimulus and demonstrated up to a 13 fold activation of the reporter gene. These experiments also suggested that there was a repressor element between –361 and –161 nucleotides of the *IL8* transcriptional start site (Fiedler et al. 1996). Subsequent studies also using upRSV examined the relative role of Activator protein-1 (AP-1), CAAT-enhancer binding protein (C/EBP) β and nuclear factor (NF)-κB sites in IL-8 production (Mastronarde et al 1998). Transfection of A549 cells with plasmid promoter constructs showed a positive cooperative effect between the (AP-1) and the NF-κB sites (Figure 5.1). Mutation of the C/EBP β site had very little effect in the presence of intact AP-1 and NF-κB sites.

Subsequent transfection studies using pRSV yielded very different results (Garofalo, 1996). Transfection of A549 cells with 5' flanking region *IL8* constructs linked to a luciferase reporter gene showed 5-fold activation over uninfected controls at 12 hours and a 4 fold induction at 24 hours. The role of different sections of the 5' flanking region of the *IL8* gene was clarified further. Casola and colleagues (2000) showed that TNF induced *IL8* expression only requires the region from –99 to –54 nucleotides from the transcriptional start site. However, this same region was insufficient for RSV-induced IL-8 production. Instead, RSV requires the AP-1 binding site at –132 to –99 nucleotides as well as an additional novel upstream element called the RSV response element (RSVRE) at -162 to -132 nucleotides. RSV infection increased RSVRE binding activities and increased synthesis of IFN regulatory factor-1 protein (IRF-1) that is present in the RSVRE complex. This finding extends the sites that cooperate with IRF-1 to include C/EBP β and AP-1.

In addition to activation elements, the 5' flanking region of *IL8* also has the capacity to silence the gene. Sequence comparison with the IFN-β promoter found a negative regulatory element (NRE) in the *IL8* promoter partially overlapping the NF-κB response element (Nourbakhsh, 2001). The NRE appears to have a dual role in *IL8* transcription as in the absence of a stimulus it appears to be involved in transcriptional silencing. However, it is required also for full induction of the *IL8* promoter in the presence of IL-1 stimulation.

None of the above studies looking at the detailed function of the 5' flanking region of *IL8* have included an investigation of the effect SNPs may have on *IL8* transcription. Yet since the publication of the original association studies concerning the *IL8* -251A polymorphism (Hull et al 2000, Hull et al 2001) that same allele has been associated with susceptibility to

tuberculosis (Ma et al 2003) enteroaggregative *E.Coli* diarrhea (Jiang et al 2003), prostate cancer (McCarron et al 2002), and greater IL-8 protein production (Jiang et al 2003). In this chapter I wished to test for function at the -251 SNP in the context of reporter constructs of the 5' flanking region of the *IL8* gene. The position of a putative repressor element between –361 and –161 nucleotides of the transcription start site of *IL8* was of particular interest to my inquiry into genetic function at the -251 polymorphism. I wished to test for -251 polymorphism function after stimulation with both TNF and RSV. Given the different mechanisms by which these agents induced *IL8* I decided to test whether -251 SNP function differed between them. In addition to looking for functional differences at this SNP with transfection studies I proposed also to test whether there were allele-specific differences in DNA-protein binding at this polymorphism.

Results

5.2.1

No evidence of a repressor element between nucleotides -389 and -169 of the 5' flanking region of the IL8 gene

Given the complexity of transcription factor binding sites in the 5' flanking region of *IL8* a reporter plasmid was constructed containing 1409 base pairs of the 5' sequence 5' to *IL8*. These constructs differed only at the –251 position (Figure 5.1) so that the -251A and -251T alleles were assessed from within normal sequence of regulatory elements. Transfection of these *IL8* constructs into A549 cells showed that maximal luciferase activity occurs 14 hours after TNF stimulation (Figure 5.2). Dr Irina Udalova's initial transfection experiments in A549 cells appeared to show increased expression of the –389 construct compared with the –169 plasmid following TNF (40 ng/ml) (Figure 5.3). However, these experiments were repeated as there was wide variability in values between experiments and there was an absence of a second co-transfection plasmid to control for transfection efficiency. As shown in Figure 5.4, there was no clear increase in expression between the –389 and the –169 plasmids after correcting for variations in transfection efficiency by co-transfection with the renilla plasmid (pRL-TK). Further experiments using pRSV and upRSV as a stimulus showed no difference in reporter gene activity between the –389 and the –169 plasmids (data not shown). No difference in expression was noted between a stimulus of TNF at 40 ng/ml and 10ng/ml (data not shown).

Figure 5.1 **Effects of the –251 polymorphism in the *IL8* 5' flanking region on reporter gene expression.** Schematic representation of up to –1409 nucleotides of the 5' flanking region of the *IL8* gene depicting the main transcription factor binding sites

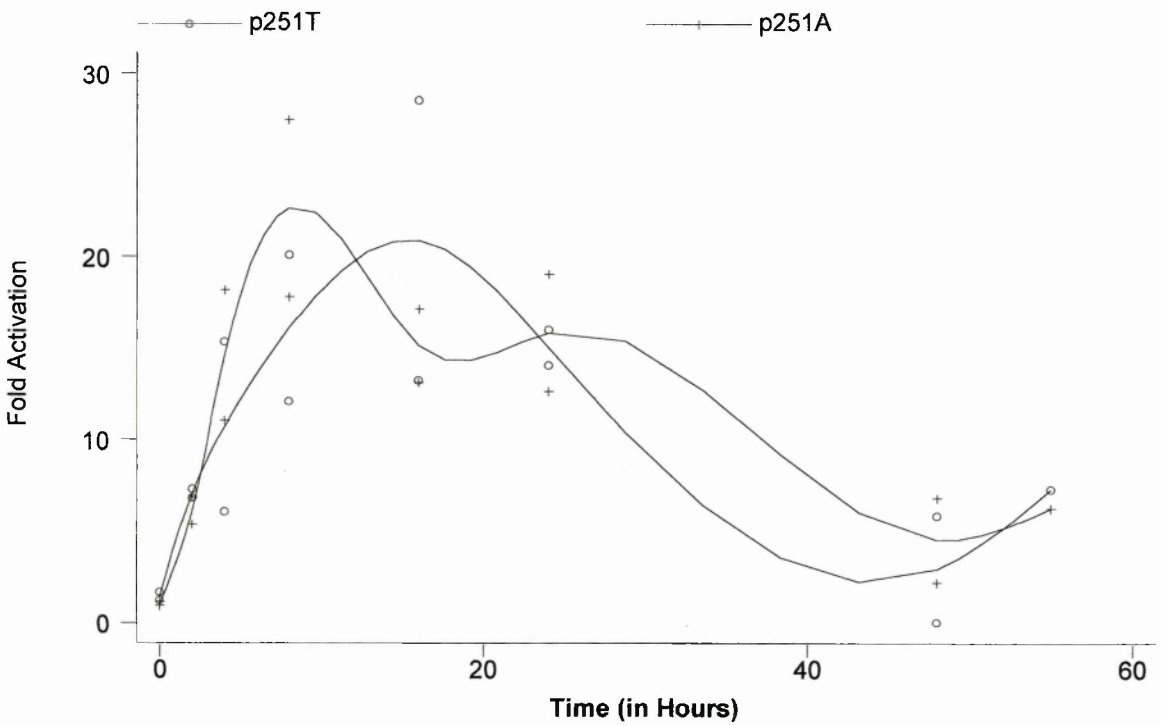


Figure 5.2 **Transcriptional activity over time of the -251A and -251T polymorphisms in the 1409 5' flanking region of the *IL8* gene in response to TNF (5 ng/ml) stimulation after transfection into A549 cells.** All data are reported as the increase observed over controls for two separate experiments. Data points are connected by the curve calculated to represent the mean fold increase.

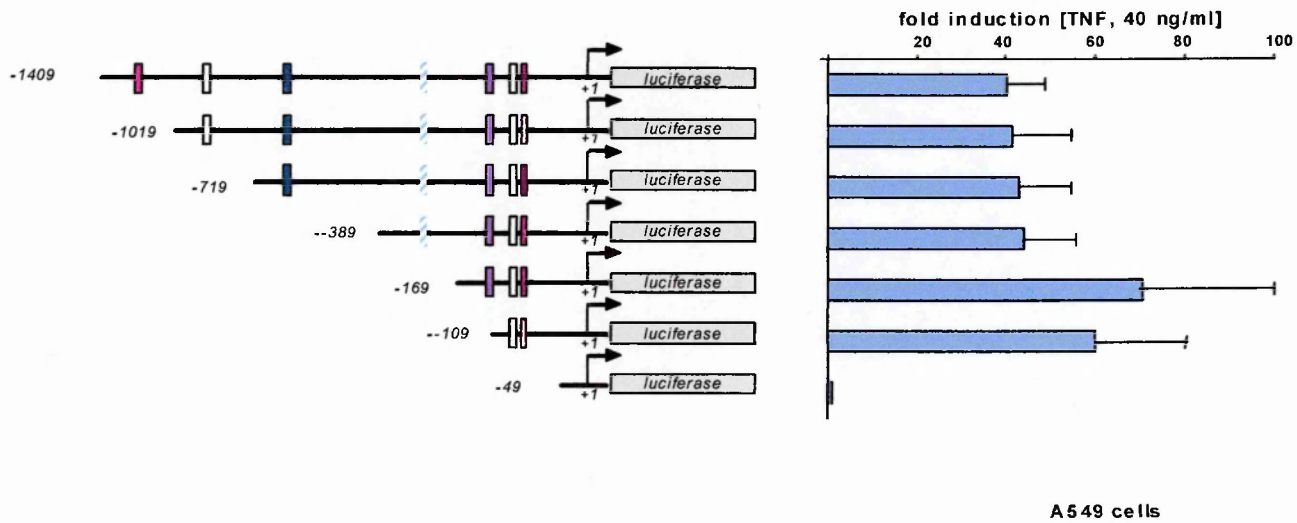


Figure 5.3.

Transcriptional activity of +1409 nucleotides of the 5' flanking region of the *IL8* gene in response to TNF (40 ng/ml) stimulation after transfection into A549 cells. All data are reported as the mean increase observed over controls for five separate experiments. Error bars represent standard error of the means. Adapted from a figure drawn by Dr Irina Udalova and used with permission of the author.

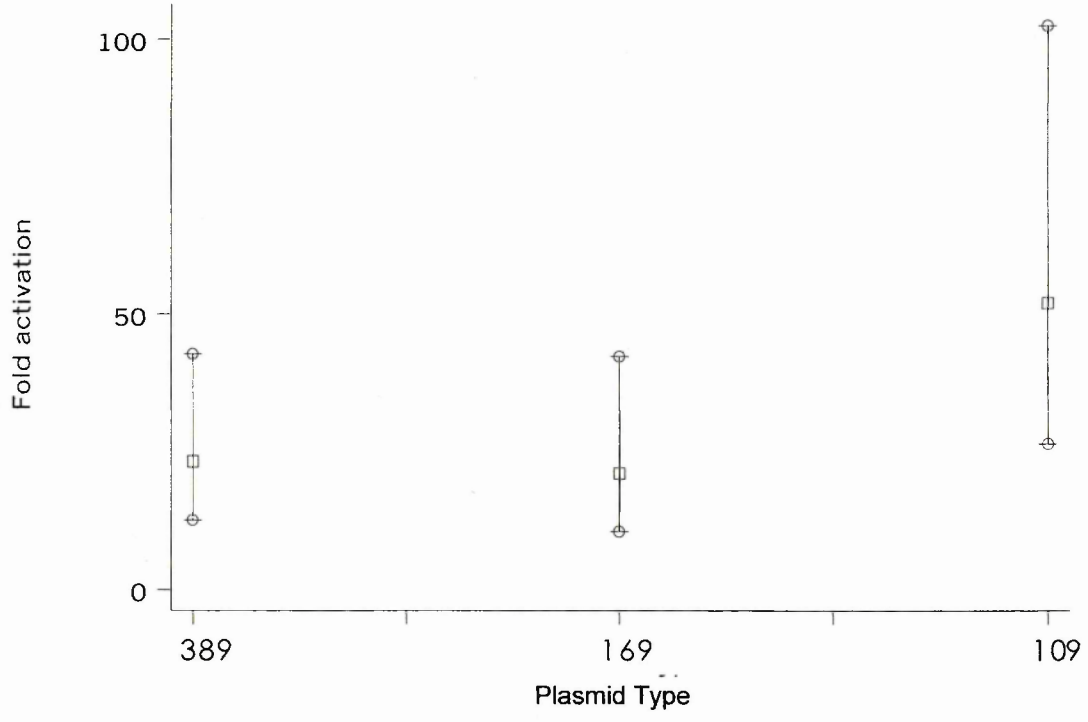


Figure 5.4

Transcriptional activity of +389 nucleotides of the 5'flanking region of the *IL8* gene in response to TNF (40 ng/ml) stimulation after transfection into A549 cells. All data are reported as the geometric mean increase in observed over controls for six observations from three separate experiments. Error bars represent the 95% confidence interval. Transfection efficiency was controlled for by co-transfection with the pRL-TK plasmid.

5.2.2

The -251 (A/T) polymorphism in response to TNF stimulation appears to be functional in H460 cells but not in A549 cells

To define whether the -251 (A/T) polymorphism had a functional effect on transcription plasmids containing either the -251A allele or the -251T allele were transfected into Human H460 pulmonary epithelial cells (H460) and Human A549 pulmonary type II epithelial cells (A549). As shown in Figure 5.5 transfection of the 1409(-251T) and 1409(-251A) constructs into both H460 and A549 cells showed significant induction in the presence of TNF. In H460 cells (Figure 5.5a) the fold rise of mean reporter gene expression in activated over non-activated cells was significantly greater for the 1409(-251T) construct (18 fold) when compared with the 1409(-251A) construct (13 fold, $p=0.008$). By comparison, transfection of these plasmids into A549 cells, as shown in Figure 5.5b, did not show a significantly different fold rise between the 1409(251T) construct (25 fold) and the 1409(251A) construct (21 fold) ($p=0.32$).

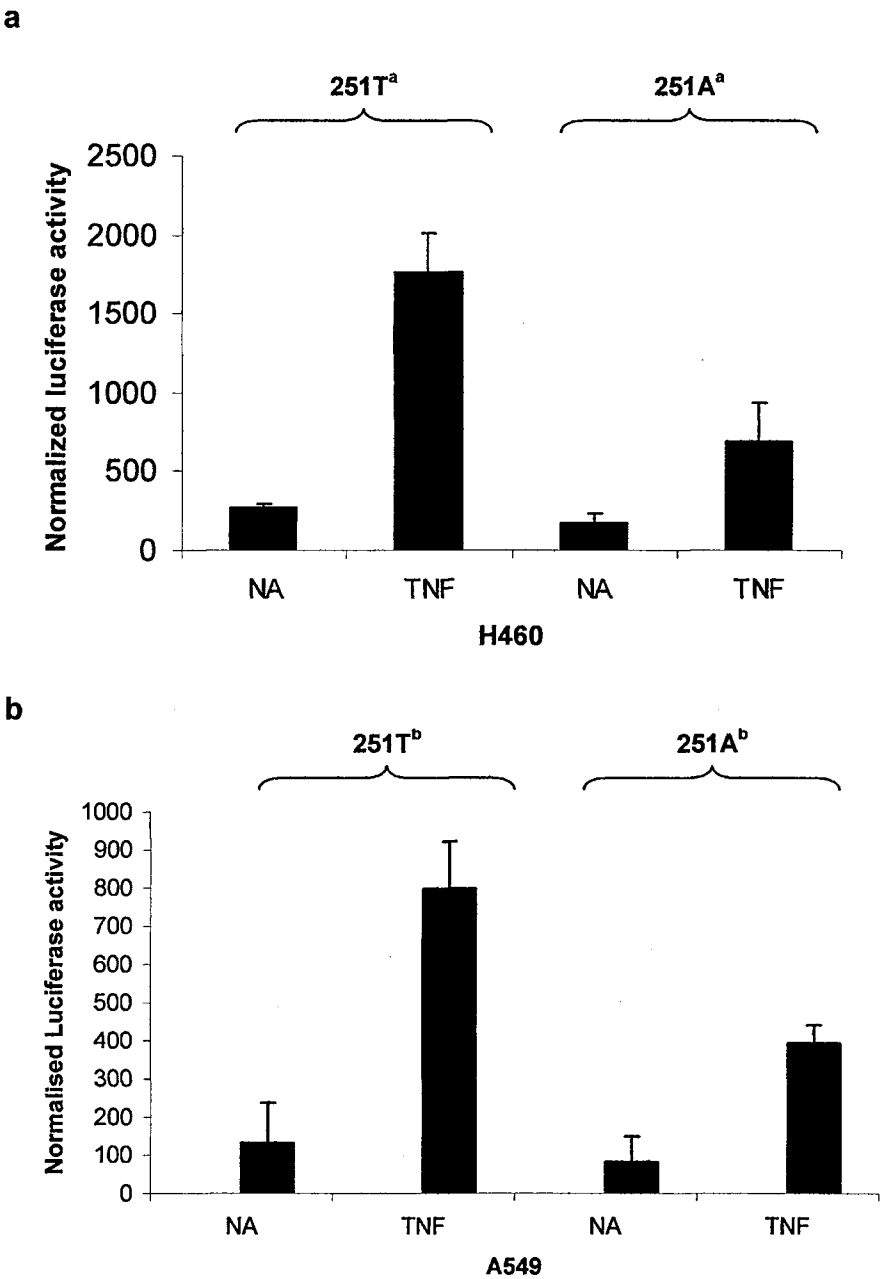


Figure 5.5 **Effects of the –251 polymorphism in the IL-8 promoter sequence on reporter gene expression after TNF stimulation.**

Figure 5.5 **Effects of the –251 polymorphism in the *IL8* promoter sequence on reporter gene expression after TNF stimulation.** H460 cells (**a**) and A549 cells (**b**) were transiently transfected with 1.4kb of the *IL8* promoter region containing either the IL8-251T or the IL8-251A allele after which the cells were stimulated with TNF (5ng/ml). Luciferase activity was corrected for transfection efficiency by co-transfection with a renilla construct. The values are expressed as the mean (± the standard error of the mean) normalized luciferase activity for non activated (NA) and activated (TNF 5ng/ml) cells from 4 independent experiments in (**a**) and 5 independent experiments in (**b**). The fold rise compares the mean normalized luciferase readings of activated over non-activated transfected cells. [a]-251T 17.9 fold rise, -251A 12.6 fold rise: paired *t* test, p=0.008; [b]-251T 25.1 fold rise, -251A 21.2 fold rise: paired *t* test , p=0.32.

5.2.3

The -251 (A/T) polymorphism in response to pRSV stimulation does not appear to be functional in either H460 or A549 cells

To define whether the -251 (A/T) polymorphism had a functional effect on *IL8* transcription following pRSV stimulation the H460 and A549 cells were again transfected with the *IL8* +1409 5' flanking region plasmid containing the two -251 polymorphic variants. A time course experiment was conducted to investigate the peak luciferase activity after the pRSV stimulus (data not shown). In both cell lines the fold activation in luciferase activity was low at between 4 and 1.5 fold and though there was much variability in the readings a time point of 14 hours after activation was selected as optimal for reporter gene activity. As shown in Figure 5.6 transfection of the 1409(-251T) and 1409(-251A) constructs into both H460 and A549 cells did not show significant induction in the presence of RSV. In H460 cells (Figure 5.6a) the fold rise of mean reporter gene expression in activated over non-activated cells was not significantly different between the 1409(-251T) construct (3.9 fold) and the 1409(-251A) construct (2 fold, p=0.17). Transfection of these plasmids into A549 cells, as shown in Figure 5.6b, did show a significantly different fold rise between the 1409(251T) construct (2.4 fold) and the 1409(251A) construct (1.7 fold) (p=0.02). However, in this series two experiments failed to show any activation on stimulation with pRSV in the presence of the -251A plasmid. When those experiments with failed activation are excluded the significant difference between the two does not persist (p=0.08). Given that there is no clear induction of the reporter gene on stimulation with pRSV in either cell line the difference in fold activation between the -251 constructs in A549 cells does not constitute adequate evidence for function at this SNP after pRSV stimulation.

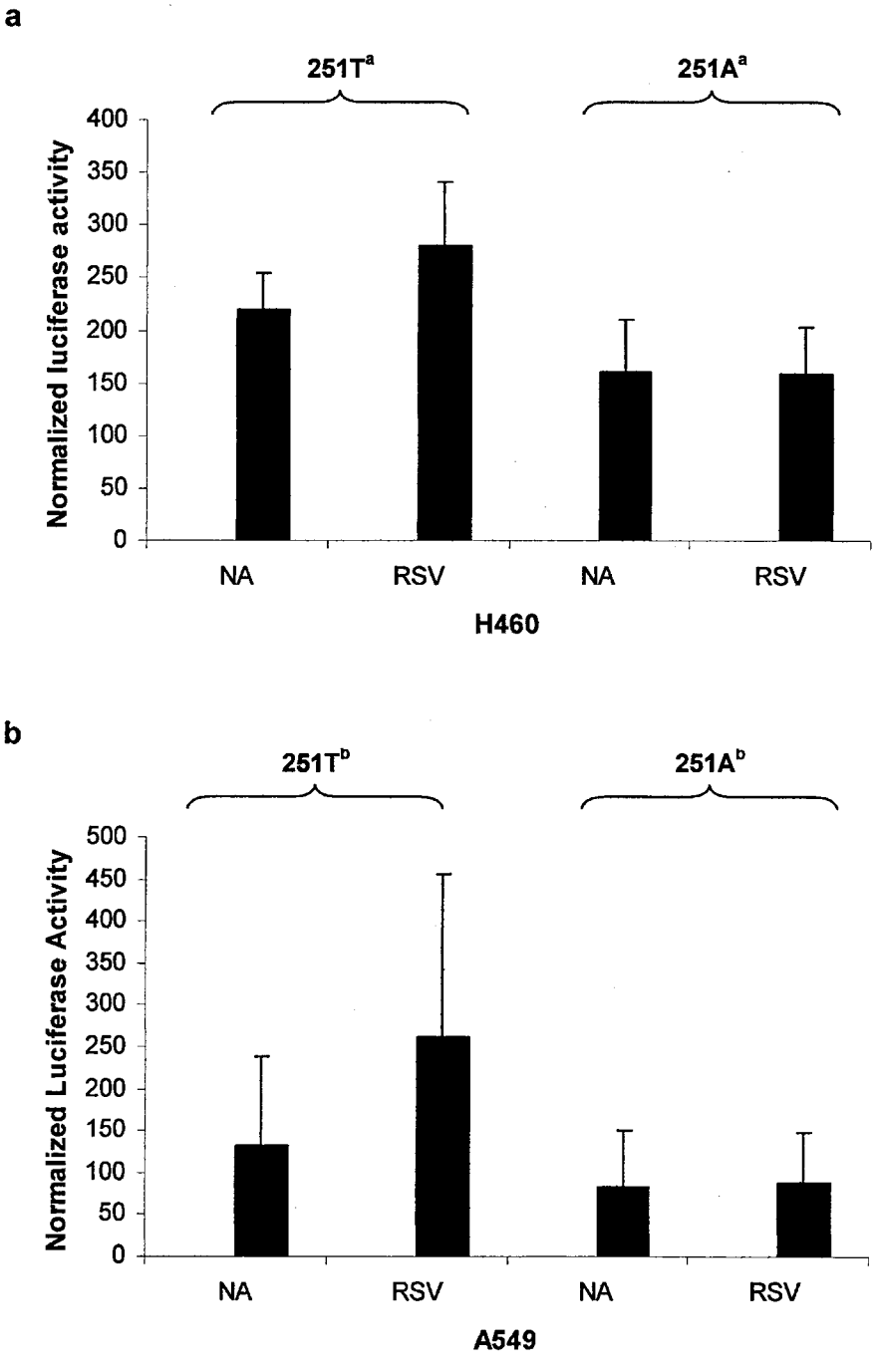


Figure 5.6 **Effects of the –251 polymorphism in the *IL8* promoter sequence on reporter gene expression after RSV stimulation.**

Figure 5.6 **Effects of the –251 polymorphism in the *IL8* promoter sequence on reporter gene expression after RSV stimulation.** H460 cells (**a**) and A549 cells (**b**) were transiently transfected with 1.4kb of the *IL8* promoter region containing either the IL8-251T or the IL8-251A allele after which the cells were stimulated with RSV (MOI 1). Luciferase activity was corrected for transfection efficiency by co-transfection with a renilla construct. The values are expressed as the mean (± the standard error of the mean) normalized luciferase activity for non activated (NA) and activated (RSV MOI 1) cells from 4 independent experiments in (**b**) and 5 independent experiments in (**c**). The fold rise compares the mean normalized luciferase readings of activated over non-activated transfected cells. [a]-251T 3.9 fold rise, -251A 2.6 fold rise: paired *t* test, p= 0.17; [b]-251T 2.4 fold rise, -251A 1.7 fold rise: paired *t* test , p= 0.02.

5.2.4

Allele-specific protein to DNA binding to the 5' flanking region of IL8 is modulated by the -251 polymorphism

We performed EMSAs to explore whether the differences in expression at the –251 SNP between the A549 and H460 cells were attributable to differential binding of nuclear proteins in the presence of that SNP. Two nucleotide probes spanning the -251A or -251T SNP were incubated in nuclear extracts from TNF stimulated H460 (Figure 5.7a) and A549 (Figure 5.7b) cells. Two specific protein-DNA complexes (designated *I* and *II*) were detected in nuclear extracts from both cells by EMSA. In H460 nuclear extracts complex *I* bound to the –251A probe but not to the –251T probe. The specificity of this interaction was defined by competition experiments in which complex *I* (**C1**) was competed only by the –251A unlabelled oligonucleotide (Figure 5.7a, *lane 3*). In A549 nuclear extracts complex *I* (**C1**) is seen also to specifically bind to the -251A probe (Figure 5.7b, *lanes 1* to *5*) but in addition it binds with lower affinity to the -251T probe. This is seen from competition which reveals that complex *I* was competed from the -251T probe by –251A unlabelled oligonucleotide more readily (*lanes 7* and *8*) than it was by the -251T unlabelled oligonucleotide (lanes 9 and 10). Complex *I* from A549 (Figure 5.8a) and H460 (Figure 5.8b) nuclear extracts was analysed by SDS-PAGE and a major band was seen at M_r~97kDa in both cases. Given that the oligonucleotide probe migrates at about 20kDa in this gel system it would suggest that the constitutive factor contains a major component of about 77kDa. Complex *II* (**C2**) binds to both the -251A and -251T probes in H460 (Figure 5.7a) and A549 (Figure 5.7b) nuclear extracts and is specifically competed. We performed additional EMSAs to investigate the observation that the consensus binding site (5'-ATGCAAAT-3') for the ubiquitous transcription factor Oct-1 (Fletcher et al. 1987) was

present on the reverse strand of the -251T allele (5'-TATCAAAT-3') but not the –251A allele (5'-TATCAATT-3'). As seen in Figure 5.9 Complex *II* appeared to contain the Oct-1 transcription factor as an EMSA using specific antibodies to Oct-1 showed a supershift of complex *II* in the presence of both the -251A and -251T probes whereas control antibodies against p65 NF-κB had no effect. No differences in binding for the –251A and –251T probes were seen between unstimulated and stimulated A549 and H460 nuclear extracts (data not shown).

a

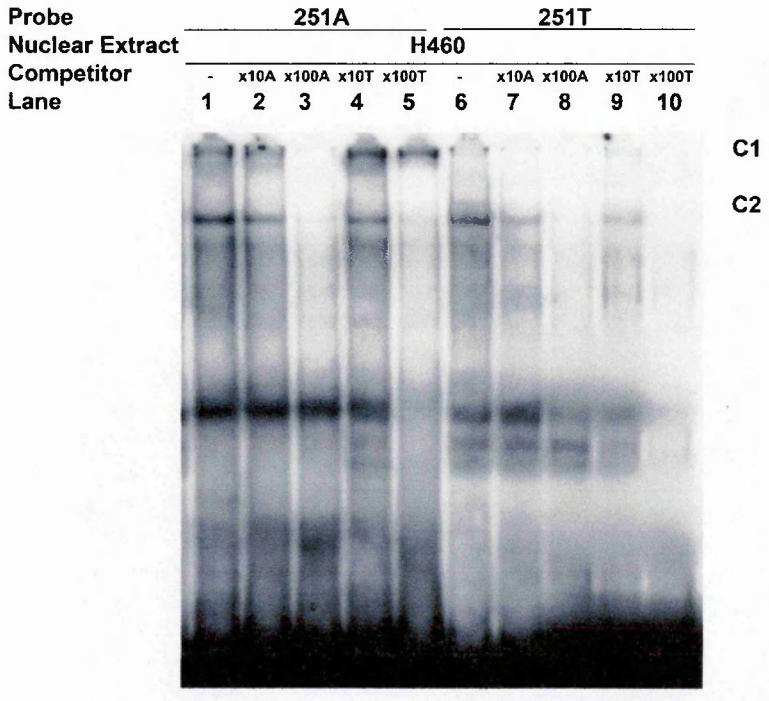


Figure 5.7 **Electrophoretic mobility assays demonstrate differences in allele-specific binding of nuclear proteins between H460 and A549 cell lines at the –251 SNP.**

b

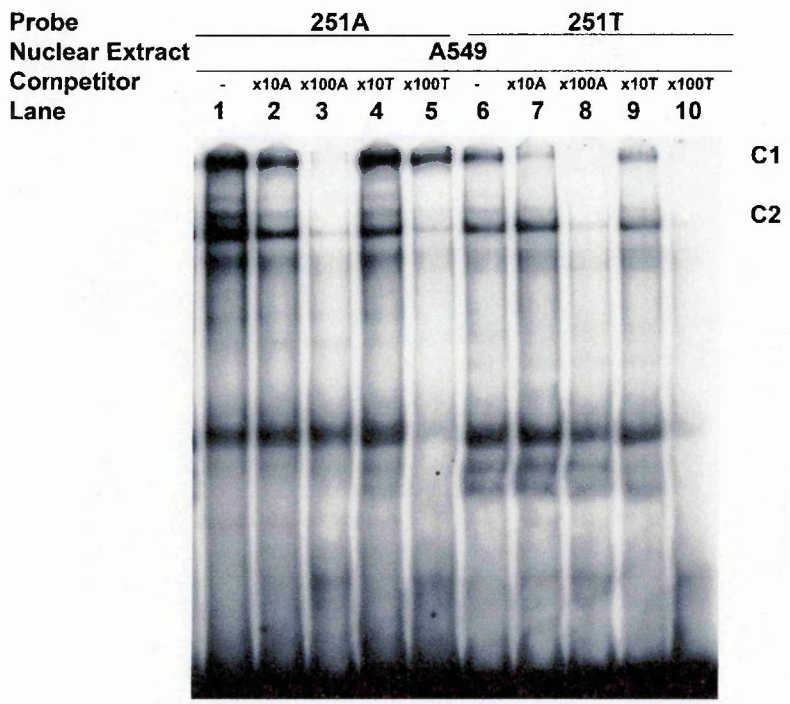


Figure 5.7 **Electrophoretic mobility assays demonstrate differences in allele-specific binding of nuclear proteins between H460 and A549 cell lines at the –251 SNP.**

Figure 5.7 **Electrophoretic mobility assays demonstrate differences in allele-specific binding of nuclear proteins between H460 and A549 cell lines at the –251 SNP.** Nuclear extracts from H460 cells (**a**) and A549 cells (**b**) harvested 6 hours after stimulation with TNF (30ng/ml) were used with probes corresponding to the -251A (*lanes 1* to *5*) and -251T alleles (*lanes 6* to *10*). Unlabelled competitor oligoduplex probes were added in molar excess (x10A and x100A refer to x10 and x100 molar excess of unlabelled -251A probe respectively, whilst x10T and x100T refer to x10 and x100 molar excess of unlabelled -251T probe respectively) as shown in *lanes 2* to *5* and *lanes 7* to *10*. Complex *I* (**C1**) specifically bound to the –251A probe in the presence of H460 (**a**) and A549 (**b**) nuclear extract. Complex *I* (**C1**) also bound to the –251T when incubated with A549 nuclear extract (**b**). Complex *II* (**C2**) specifically bound to the –251A and to the -251T probes in the presence of H460 (**a**) and A549 (**b**) nuclear extract and was competed by x100 molar excess of unlabelled -251A probe (*lanes 3* and *8*) as well as x100 molar excess of unlabelled -251T (*lanes 5* and *10*).

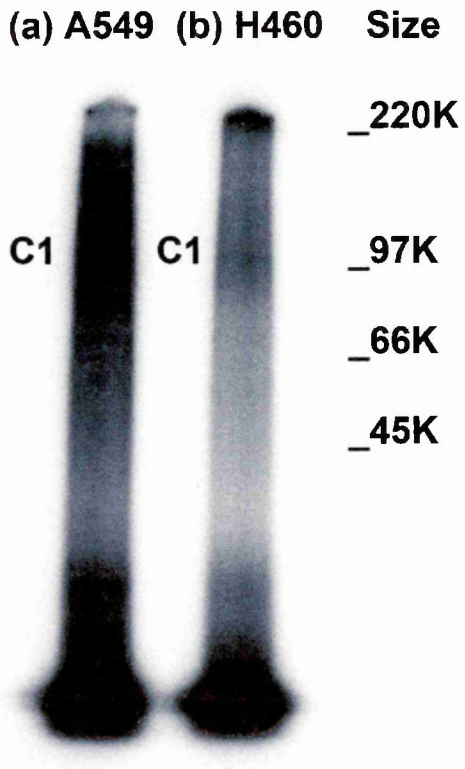


Figure 5.8 **UV cross linked products of a bromodeoxyuridine-substituted probe spanning the –251A SNP with nuclear extracts from H460 and A549 cells**. Complexes were separated by EMSA, subjected to UV illumination, cut from the gel, and analyzed on 10% Bis-Tris MOPS-SDS NuPAGE gels. Rainbow full range molecular marker (Amersham Pharmacia Biotech) is indicated on the right. **a**, the complex *I* (**C1**)taken from A549 cells is seen at about ~97 kDa. **b**, complex *I* (**C1**) taken from H460 cells is also seen at about ~97kDa.

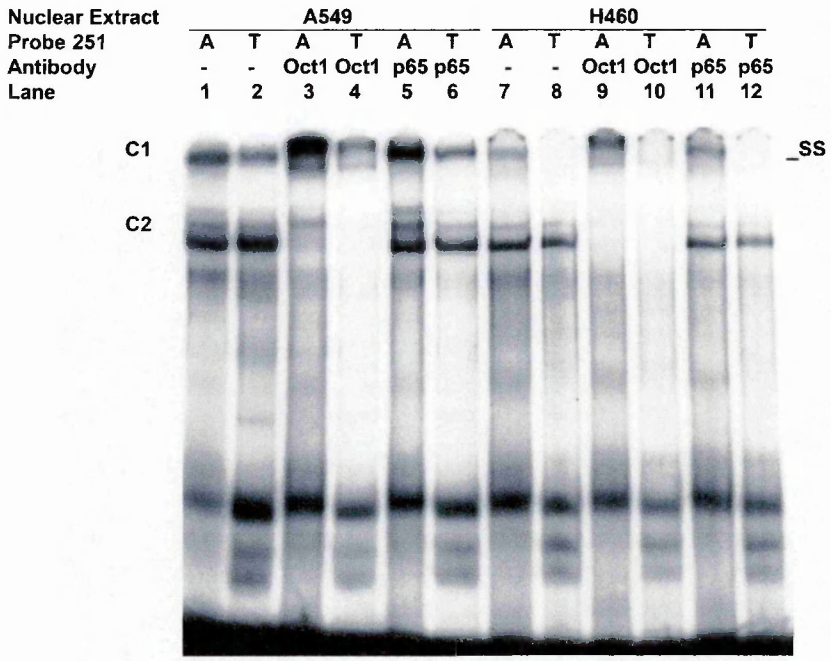


Figure 5.9 **Electrophoretic mobility shift assays demonstrate that Oct-1 can bind to the -251A allele and to the -251T allele**. EMSA showing nuclear extract from A549 cells (*lanes 1* to *6*) and H460 cells (*lanes 7* to *12*) were incubated with probe in the absence (*lanes 1, 2, 7* and *8*) or presence of antibodies (*lanes 3, 4, 9* and *10*) to Oct-1. Complex *I* (**C1**) is present after incubation with the -251A and -251T probes in A549 nuclear extract (*lane 1* and *2*) and after incubation with the -251A probe in H460 nuclear extract (*lane 7*). Complex *II* (**C2**) is present after incubation with the -251A and -251T probes in both A549 nuclear extract (*lane 1* and *2*) and in H460 nuclear extract (*lane 7* and *8*). A supershifted band (indicated by SS_) formed with the anti-Oct-1 antibody is seen in the presence of the -251A and -251T probes in both A549 (*lanes 3* and *4*) and H460 (*lanes 9* and *10*) nuclear extract. No supershifted bands were seen in the presence of the control antibody to p65 NF-κB.

Discussion

5.3.1

IL8 transcriptional regulation by a putative repressor element in the 5' flanking region of the gene

In common with reports from another laboratory (Casola et al. 2000) I could not reproduce the result reported by Fiedler et al. (1996) showing a putative repressor element between –361 and –161 nt 5' of the *IL8* gene in response to TNF, upRSV or pRSV stimulation. One possible explanation for this might lie in the non-standardised nature of the upRSV preparation as it may contain a number of cytokines at concentrations which vary between laboratories. For instance, two different upRSV preparations have been shown to produce very different results in stimulating NO expression (Tsutsumi et al. 1999; Kao et al. 2001).

5.3.2

Transcriptional activity of the 251 (A/T) promoter polymorphism of IL8 in response to TNF stimulation

Respiratory epithelial cells initiate and regulate the innate immune response (Garofalo and Haeberle, 2000; Garofalo et al. 2001) to RSV infection. In addition to the differences seen in *IL8* regulation between primary and immortalized respiratory cells (Chpater 4) I have shown that there may also be differences in regulation between immortalized cell lines. Consistent with the increased *in vivo* allele specific expression of haplotype 1 in H460 cells I noted significantly increased reporter gene expression associated with the -251T polymorphism in these same cells. Though the pattern of reporter gene activation was similar in both cell lines the –251 SNP did not appear to be functional when the *IL8* constructs were transfected into A549 cells. Also, I describe differential binding of nuclear

factors between the –251A and –251T polymorphisms in the H460 cells which were not present in A549 cells. Complex *I*, which contained an unidentified nuclear factor of about 77kDa, bound in the presence of the -251A allele alone in H460 cells. However, complex *I* bound in the presence of both -251 alleles in A549 cells. A second complex (complex *II*) was detected in the presence of both -251 alleles. However, there is no difference in the binding of complex *II* between the two alleles or between the cell lines. It is therefore unlikely that either complex *II*, or the factor Oct-1 which was detected within it, play a role in explaining the differences in function at the -251 polymorphism.

Previously differences in the functional regulation of the same haplotype, in this case IL-6, between immortalized cell lines have been described (Terry et al 2000). In addition, there is evidence that *IL8* is regulated in a different manner between H292 lung epithelial cells and A549 cells (Roger et al 1998, Newton et al 1996). This discrepancy in the regulation of *IL8* between respiratory cell lines has not previously been associated with functional polymorphisms.

The extent to which one believes the -251 polymorphism is functional rests in part on which of the two respiratory cell lines best represents the gene regulation seen in primary respiratory cells. H460 cells are derived from a culture of a Large Cell lung carcinoma and have hypotriploid karyotype with a modal chromosomal number of 58. A549 cells are derived from a Small Cell carcinoma and have hypotriploid karyotype as well but differ in that they have a modal chromosomal number of 12 (Giard et al. 1973). The A549 cell was selected for the low levels of *IL8* that are secreted in the inactivated state. However, examination of a number of other respiratory epithelial cell lines by the Kwiatkowski

laboratory, including H460 cells, has shown that moderate basal expression of *IL8* is more common. This implies that the patterns of *IL8* induction in A549 cells may be unusual compared with similar cell lines. However, it seems that *IL8* expression in A549 cells is closer to that seen in primary cells. On these grounds the A549 cells appear to be a better model within which to study *IL8* functional genetic regulation. However, in the absence of transient transfection experiments in primary respiratory cells the question of function at the -251 polymorphism would be difficult to determine definitively.

Comparing the results of the transfection experiments with the genetic epidemiological data may also help us define whether the -251 polymorphism is functional. The *IL8* -251A allele has been associated with susceptibility to bronchiolitis (Hull et al 2000), tuberculosis (Ma et al 2003) enteroaggregative *E.Coli* diarrhea (Jiang et al 2003), prostate cancer (McCarron et al 2002), and greater *IL8* protein production (Jiang et al 2003). Hull and co-workers reported that the association between the *IL8* -251A allele and RSV-induced bronchioilitis was further refined by haplotype analysis. They suggested that the functional allele might lie on chromosomes of haplotype 2 and that the -251A may not be a functional SNP (Hull et al 2001). However, the association with haplotype 2 was not reproduced in the context of these other disease states though this may have been due to the smaller sample size used in these studies (Ma et al 2003, Jiang et al 2003). The data from transfection experiments in A549 cells is consistent with the most detailed genetic epidemiological data (Hull et al. 2001) that suggests that the -251 SNP is not functional. Moreover, the transfection studies in A549 cells do not contradict the results of haplotype-specific experiments in primary cells. Therefore taking the data on haplotype-specific expression, genetic epidemiology and *IL8*

induction together it is likely that the A549 cell is a more accurate model for *IL8* regulation and that the -251 SNP is not functional after TNF stimulation.

I have not conducted further experiments looking to see if any potential functional genetic effect at the 251 polymorphism could be seen when it was placed within the shorter *IL8* promoter constructs of 1019, 789, or 389 nucleotides respectively. It has been shown, by the Kwiatkowski laboratory (Figure 5.3) and others (Garofalo et al. 1996), that the shorter constructs have the same inducibility in response to TNF as the longer constructs. Moreover, an examination of the -251 SNP within the context of the whole promoter is more likely to yield data that is free from experimental artifacts.

5.3.3

Transcriptional activity of the 251 (A/T) promoter polymorphism of IL8 in response to RSV stimulation

The luciferase activity of the pRSV stimulated cells in these experiments was variable and below the 4 to 5 fold levels of activation previously reported (Casola et al. 2000). In H460 cells there was no evidence of function at the -251 SNP in response to pRSV stimulation. In A549 cells the pRSV failed to induce any activation of the reporter gene of the -251A plasmid in some experiments. This produced data showing a statistically significant difference between the two plasmids containing the -251 SNPs. However, the levels of reporter gene activation were not sufficiently distinct from background levels to allow us to have confidence in this result. However, experiments where pRSV was used in combination with a known concentration of a cytokine, such as IL-1 or TNF, may produce sufficient

synergistic activation of *IL8* constructs. This might allow further examination of function at the -251 SNP in circumstances which are closer to that seen in RSV infections *in vivo*.

5.3.4

The limitations of transfection experiments in examining the functional genetics of SNPs

The transfection experiments described here examine the effect of a single SNP within a section of the 5' flanking region of the *IL8* gene that is engineered into bacterial plasmid and then expressed within an immortal respiratory epithelial cell line. In addition to the problems of using immortalized cell lines that have been discussed above there are a number of other limitations to the transient transfection experimental strategy.

Transient transfection assays have a peak of reporter gene activity which occurs at 34 hours after transfection and 14 hours after cell activation. However, we know from studies of *IL8* transcription in A549 cells (Figure 3.4) that whilst the peak of *IL8* induction after TNF stimulation occurs at 16 hours the peak of *IL8* induction after RSV stimulation occurs as late as 36 hours. One of the reasons that the reporter gene induction after RSV stimulation in these transfection experiments was so poor could be that the peak transient transfection efficiency did not coincide with the peak of *IL8* induction. Previous experiments (data not shown) have attempted to address this by reversing the order of activation and transfection such that cells are transfected 12 hours after activation. This resulted in some increase in reporter gene induction but no difference in activation between -251 SNPs. However, the efficiency of plasmid transfection in activated cells is much reduced (Irina Udalova, personal communication) which calls into question the validity of this unorthodox approach.

Testing for the function at the remaining SNPs is complicated by both their position in intronic and untranslated regions of the *IL8* gene (Figure 1.6) and their location beyond the translational start of *IL8* that occurs in exon 1. This latter factor prevents the design of a construct where the remaining SNPs can be assessed in the context of the normal sequence of regulatory elements found in the promoter. To overcome this technical limitation EMSA studies were used to test for allele-specific protein-DNA binding.

5.3.5

Concluding Points

- There is no evidence of a repressor element between nucleotides -389 and -169 of the 5' flanking region of the *IL8* gene.
- The -251 (A/T) polymorphism in response to TNF stimulation appears to be functional in H460 cells but not in A549 cells.
- The -251 (A/T) polymorphism in response to pRSV stimulation does not appear to be functional in either H460 or A549 cells.
- Allele-specific protein to DNA binding to the 5' flanking region of *IL8* is modulated by the –251A polymorphism when using nuclear extracts from H460 and A549 cell lines.
- A constituent of the allele-specific binding complex at the –251A polymorphism contains a major component of about 77kDa.
- The data from the haplotype-specific expression studies and the genetic epidemiology analysis taken together suggest that the A549 cell is a more accurate model for *IL8* regulation in primary respiratory epithelial cells. This implies that the -251 SNP does not make a functional contribution to haplotype-specific expression *in vivo*.
- Testing for the function at the remaining SNPs with transfection experiments is complicated by both their position in intronic and untranslated regions of the *IL8* gene and their location beyond the translational start of *IL8* that occurs in exon 1. To overcome this technical limitation EMSA studies were used to test for allele-specific protein-DNA binding.

CHAPTER SIX

EXAMINATION OF NUCLEAR FACTOR BINDING TO SNPs IN THE *IL8* GENE

6.1.1

Introduction

Electrophoretic mobility shift assays (EMSA) were used to test whether the observed differences in primary respiratory cell haplotype-specific expression of *IL8* could be due to differential binding of transcription factors at SNPs within the *IL8* haplotypes. Studies of the -251 polymorphism using transient transfection experiments had failed to produce clear evidence of function at this SNP. Testing for the function at the remaining SNPs with transient transfection experiments was complicated by both their position in intronic and untranslated regions of the *IL8* gene (Figure 1.6) and their location beyond the translational start of *IL8* which occurs in exon 1. To overcome this technical limitation EMSA studies were used to test for allele-specific protein-DNA binding. In the absence of clear genetic epidemiological data to implicate a single SNP we sought to screen five SNPs in the *IL8* gene (Intron 1: +396 G/T; Intron 1: +781 T/C; Intron 2: +1238 InsA; Intron 3: +1633 T/C and 3'end terminal: +2767 T/A) for DNA to protein interactions using EMSAs. As I have discussed in Chapter 5 there is evidence to suggest that the H460 cell line may not be a good model for *IL8* gene regulation in primary respiratory cells. EMSA experiments in this chapter are therefore restricted mainly to using nuclear extracts from A549 cells. Where a specific transcription factor of interest was required as a positive control nuclear extracts from other cell lines were utilized. In the course of these EMSA experiments I wished to

Figure 6.1 **Screening for differences in allele-specific binding of nuclear proteins at *IL8* polymorphisms using EMSAs with nuclear extracts from A549 cells after incubation with TNF (a) or RSV (b).**

determine whether TNF stimulation or RSV infection of A549 cells resulted in different patterns of allele-specific DNA-protein binding.

Results

6.2.1

Screening for allele-specific binding of nuclear factors to SNPs within the IL8 gene

EMSAs used nuclear extracts from A549 cells stimulated with TNF and probes corresponding to each allele of the +396, +781, +1238, +1633, and +2767 SNPs to screen for allele-specific differences in binding. As shown in Figure 6.1a significant differential binding was observed only at the +781 polymorphism. This differential binding at the +781 SNP was confirmed using nuclear extract from A549 cells stimulated with RSV (Figure 6.1b).

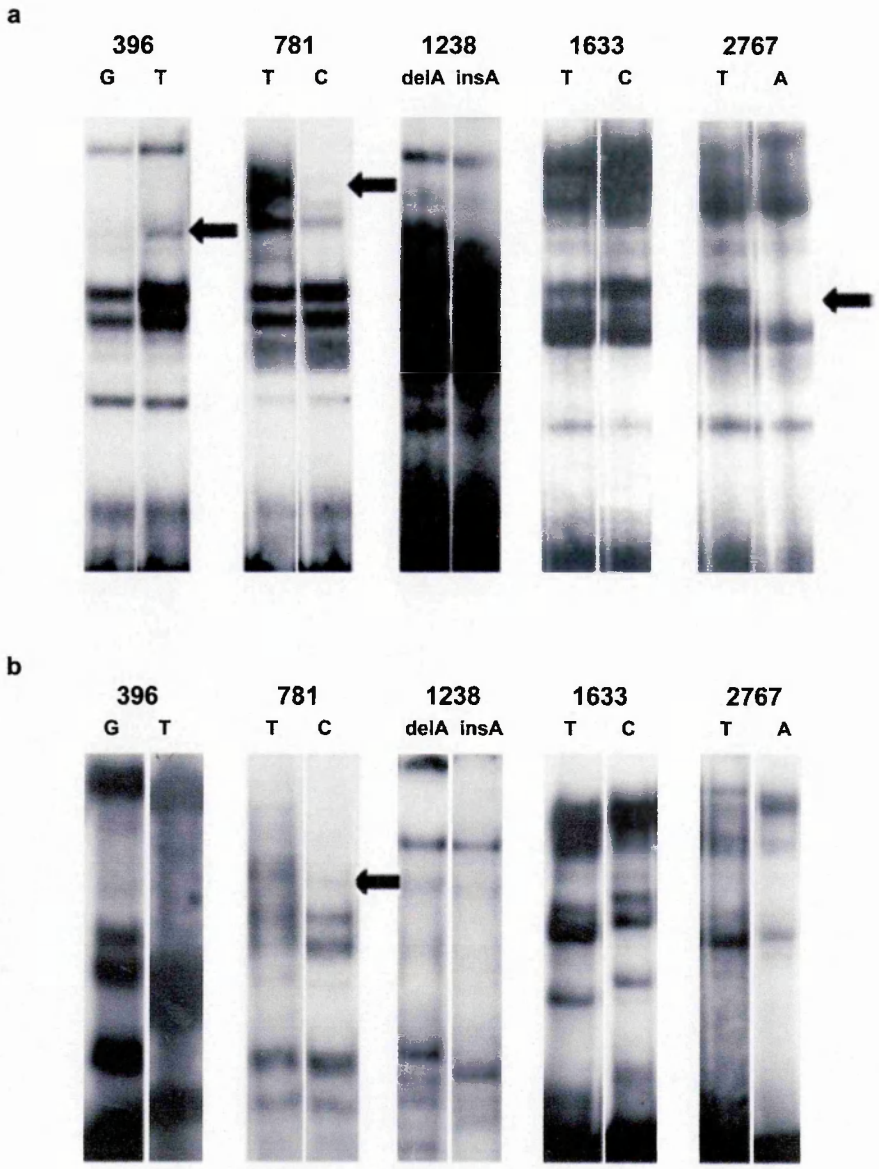


Figure 6.1 Screening for differences in allele-specific binding of nuclear proteins at *IL8* polymorphisms using EMSAs with nuclear extracts from A549 cells after incubation with TNF (**a**) or RSV (**b**).

Figure 6.1

Screening for differences in allele-specific binding of nuclear proteins at *IL8* polymorphisms using EMSAs with nuclear extracts from A549 cells after incubation with TNF (a) or RSV (b). (**a**) When incubated with nuclear extracts from A549 cells no significant allele specific differences in binding could be seen at the +1238 polymorphism (*lanes 5* and *6*) and the +1633 polymorphism (*lanes 7* and *8*). Allele-specific differences in binding were seen at the +396 polymorphism (*lanes 1* and *2*) +781 polymorphism (*lanes 3* and *4*) and the +2767 polymorphism (*lanes 9* and *10*) (indicated by the arrows). The specificity of the differential binding seen at the +396 and the +2767 polymorphisms was not confirmed in subsequent competition experiments (data not shown). (**b**) Differential binding at the +781 SNP was confirmed by EMSA studies using nuclear extracts from A549 cells stimulated with RSV (indicated by the arrow). No significant differential binding was seen for the +396, +1238, +1633 and the +2767 polymorphisms in nuclear extracts from RSV stimulated cells.

The +2767 T/A probe sequence contains a high proportion of adenine (A) and thymidine (T) bases. EMSA competition with either the poly(dI).poly(dC) non-specific inhibitor or the poly(dA).poly(dT) non-specific inhibitor yielded equivocal results (data not shown). Optimisation experiments suggested that non-specific inhibition with both poly(dI).poly(dC) and poly(dA).poly(dT) (Figure 6.2) gave the best complex resolution.

Initial EMSA experiments showed some binding of nuclear factors to the +396 and the +2767 polymorphisms (Figure 6.1a) that on further testing using competition proved to be non-specific and were not examined further. No significant differential binding was seen for the remaining SNPs (+396, +1238, +1633 and +2767) in nuclear extracts from RSV stimulated cells.

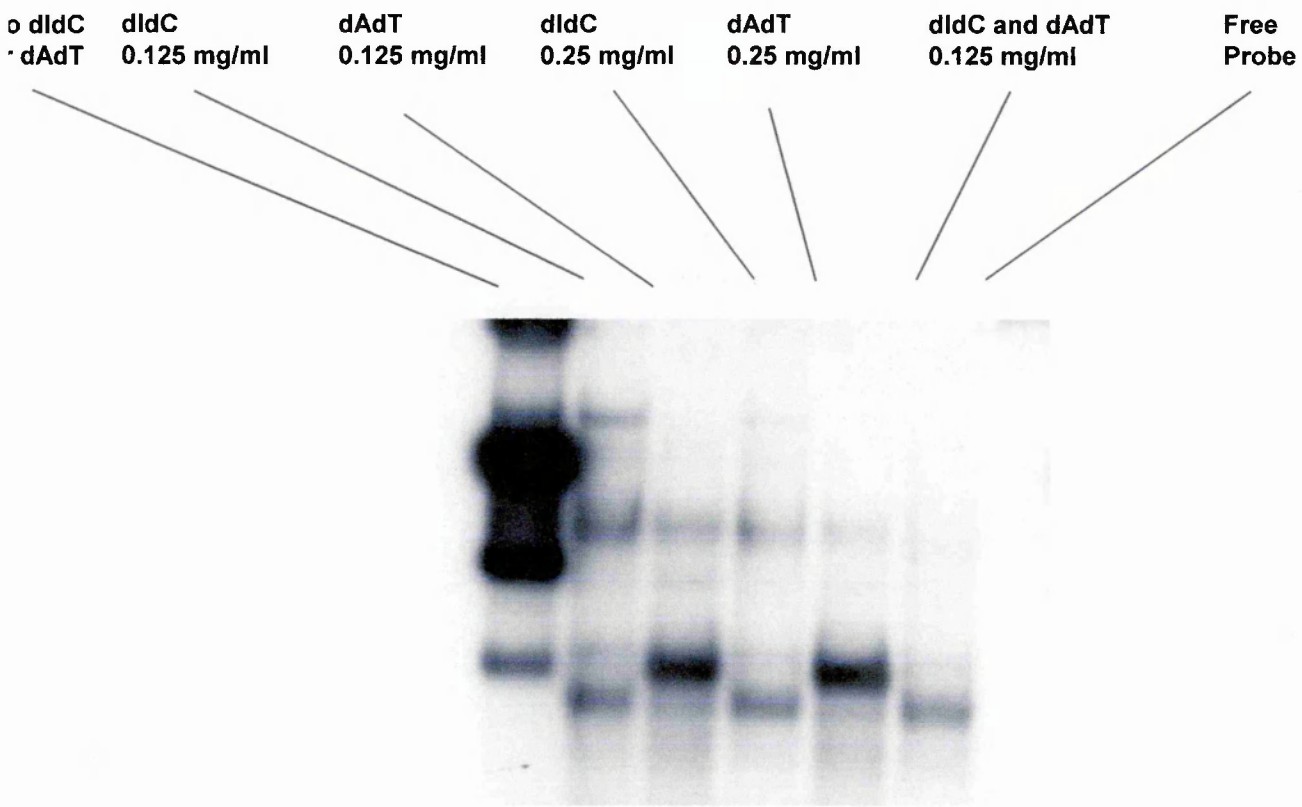


Figure 6.2

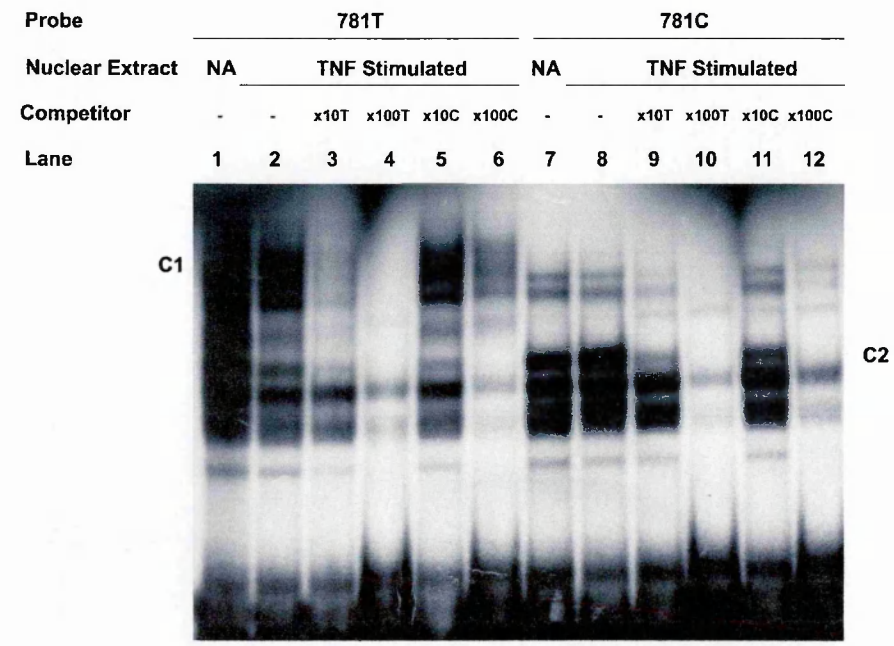
Optimisation of EMSA of *IL8* 2767 SNP binding complexes in response to RSV infection of A549 cells. Nuclear extracts prepared from RSV-infected cells after 16 hours of incubation and used for EMSA. The 2767 radiolabeled probe contains a high number of adenine (A) and Thymidine (T) bases. In order to optimize the binding conditions with the correct type and concentration of non-specific inhibitor poly(dI).poly(dC) and poly(dA).poly(dT) were incubated alone at standard (0.125 mg/ml) and twice standard (0.25 mg/ml) concentrations as well as together in the binding reaction. Optimal resolution of binding was achieved with both poly(dI).poly(dC) at 0.125 mg/ml and poly(dA).poly(dT) at 0.125 mg/ml.

6.2.2

Confirmation of allele-specific binding at the +781 SNP by competition

The specificity of binding to the +781 SNP is shown by competition experiments (Figure 6.3). Here, complex *I* was seen to bind in a constitutive manner in the presence of the +781T allele with non-activated, TNF-stimulated nuclear extract (Figure 6.3a, *lanes 1* and *2*) and RSV-stimulated nuclear extract (Figure 6.3b, *lane 1* and *3*). Complex *I* was competed only by the +781T unlabelled oligonucleotide in both TNF-stimulated (Figure 6.3a, *lanes 3* to *4*) and RSV-stimulated nuclear extract (Figure 6.3b, *lanes 4* and *5*). A second complex, complex *II*, bound constitutively to the +781C allele in the presence of TNF-stimulated nuclear extract (Figure 6.3a, *lanes 7* and *8*) at a greater affinity than the +781T allele (Figure 6.3a, *lane 1* and *2*). As the degree of differential binding between alleles was less distinct this complex was not studied further.

a



b

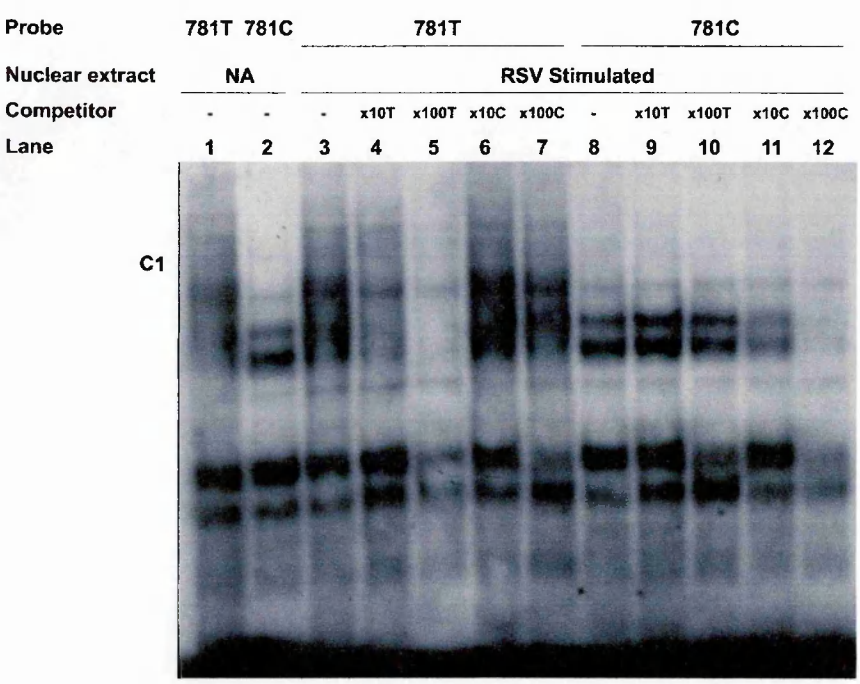


Figure 6.3 **Allele-specific binding of nuclear proteins at position +781.**

Figure 6.3

Allele-specific binding of nuclear proteins at position +781. (**a**) Probes were used corresponding to the +781T allele (*lanes 1* to *6*) and +781C allele (*lanes 7* to *12*) in EMSA with nuclear extracts from A549 cells harvested after no activation (NA) and after stimulation with TNF. Unlabelled competitor oligoduplex probes were added in molar excess as shown in *lanes 3* to *6* (x10T and x100T refer to x10 and x100 molar excess of unlabelled 781T probe respectively) and *lanes 9* to *12* (x10C and x100C refer to x10 and x100 molar excess of unlabelled 781C probe respectively). Complex *I* (C1) bound specifically to the +781T probe. Complex *II* (C2) bound to the +781C probe (*lanes 7* and *8*) with greater affinity than the +781T probe (*lanes 1* and *2*). (**b**) Probes were used corresponding to the +781T allele (*lanes 1* and *lanes 3* to *7*) and +781C allele (*lane 2* and *lanes 8* to *12*) in EMSA with nuclear extracts from A549 cells harvested after no activation (NA) and after stimulation with RSV. Unlabelled competitor oligoduplex probes were added in molar excess as shown in *lanes 4* to *7* (x10T and x100T refer to x10 and x100 molar excess of unlabelled 781T probe respectively) and *lanes 9* to *12* (x10C and x100C refer to x10 and x100 molar excess of unlabelled 781C probe respectively). Complex *I* (C1) bound specifically to the +781T probe.

6.2.3

c-Myb was not demonstrated to be a part of the complexes binding in the presence of the +781T SNP

We searched for potential transcription factor binding sites in the +781T and +781C probes using a pattern search program in MATINSPECTOR (Quandt et al. 1995) that uses the TRANSFAC database (Wingender et al. 2000). Only in the +781T probe did we find a potential transcription factor binding site, in this case c-Myb, that was not present in the +781C probe. We found no transcription binding sites that were present in the +781C probe but absent in the +781T probe. The c-Myb consensus binding site, 5'-YAACNG-3' (Y and N indicate a pyrimidine and any base, respectively) has a central section (5'-YAAC-3') that is essential for binding. The +781T lies in the centre of this essential transcription factor binding section (R: 5'-tgaaC**A**ACtcctataag-3'; R indicates the reverse strand whilst the SNP is marked in bold). The c-Myb site is disrupted therefore by the change from +781T (R: 5'-tgaaC**A**ACtcctataag-3') to +781C (R: 5'-tgaaC**G**ACtcctataag-3').

In order to determine whether c-Myb was expressed in A549 cells a western blot was performed on nuclear extract from non-activated, TNF stimulated and RSV infected A549 cell nuclear extracts (Figure 6.4). C-Myb was not detected in A549 cell nuclear extract despite TNF or RSV stimulation. By contrast c-Myb was strongly expressed in Jurkat whole cell lysate, non-activated Jurkat cell nuclear extract and PMA stimulated Jurkat cell nuclear extract.

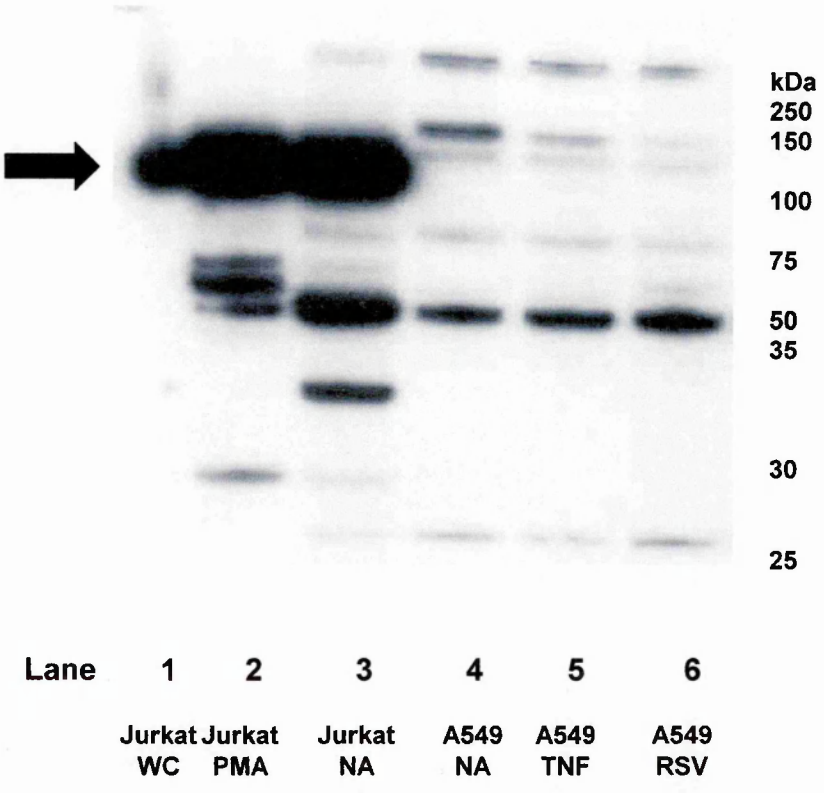


Figure 6.4

Western Blot of c-Myb protein in Jurkat cell lysate, Jurkat cell nuclear extract and A549 cells nuclear extract. When assayed using anti-c-Myb antibody c-Myb protein (arrow) was detected in Jurkat whole cell extract (**WC**) (*lane 1*), nuclear extract stimulated with PMA (**PMA**) (*lane 2*) and non-activated (**NA**) nuclear extract (*lane 3*). No c-Myb protein was detected in A549 nuclear extract derived from non-activated (**NA**) (*lane 4*), TNF (**TNF**) stimulated (*lane 5*) and RSV (**RSV**) stimulated (*lane 6*) nuclear extracts. Jurkat cells were activated with PMA at 50 ng/ml for 20 hours. A549 cells were infected with RSV at an MOI of 1 and incubated for 24 hours. A549 cells were stimulated with TNF at 5 ng/ml for 24 hours. Equal amounts of protein from control and activated cells were assayed using anti-c-Myb clone 1-1 antibody. The molecular mass (in kDa), determined by co-electrophoresed standards, is indicated on the right.

The absence of the c-Myb protein from A549 nuclear extract precludes the transcription factor being part of the allele-specific binding complex (C1) noted to be binding in the presence of the +781T SNP (Figure 6.3). However, I wished to test whether the TRANSFAC predictions of binding were correct and wondered whether c-Myb may play a role in *IL8* gene regulation in primary lymphocytes. To test these hypotheses I conducted studies of DNA-protein interaction using nuclear extract derived from Jurkat cells. In this cell line binding appeared to be less specific for both the +781T and +781C probes as C1 binding to the +781T probe was competed by 100 fold molar excess of unlabelled +781C probe whilst C2 was competed by 100 fold molar excess of the unlabelled +781T probe (data not shown).

In contrast, as shown in Figure 6.5a, there was avid (*lane 1*) and specific binding (*lane 3*) to a probe containing the c-Myb consensus sequence (C1) but not to the probe containing the mutated c-Myb sequence (*lane 6*). Furthermore, as demonstrated in Figure 6.5b, supershift bands were observed when the c-Myb consensus sequence probe was incubated in the presence of two different anti-c-Myb antibodies (*lane 5* and *lane 9*). No supershift bands were seen in the presence of anti-c-Myb antibodies and the mutated c-Myb consensus sequence probe, the +781T probe or the +781C probe. Incubation of the c-Myb consensus probe and mutated c-Myb consensus sequence probe with non-specific antibody (anti-NFκB p65) resulted in no supershift bands (data not shown).

As two other members of the Myb family of transcription factors, A-Myb and B-Myb, share the same DNA binding domain (DBD) I sought to test whether the ubiquitously expressed B-Myb may be binding to the +781T probe. However, incubation of either the +781T or the +781C probe with anti B-Myb antibodies did not result in supershift bands (data not shown). From these data I can conclude that c-Myb is present in Jurkat nuclear extract and does bind to the c-Myb consensus probe. However, neither the c-Myb nor B-Myb proteins bind in the presence of the +781T or the +781C SNPs *in vitro*.

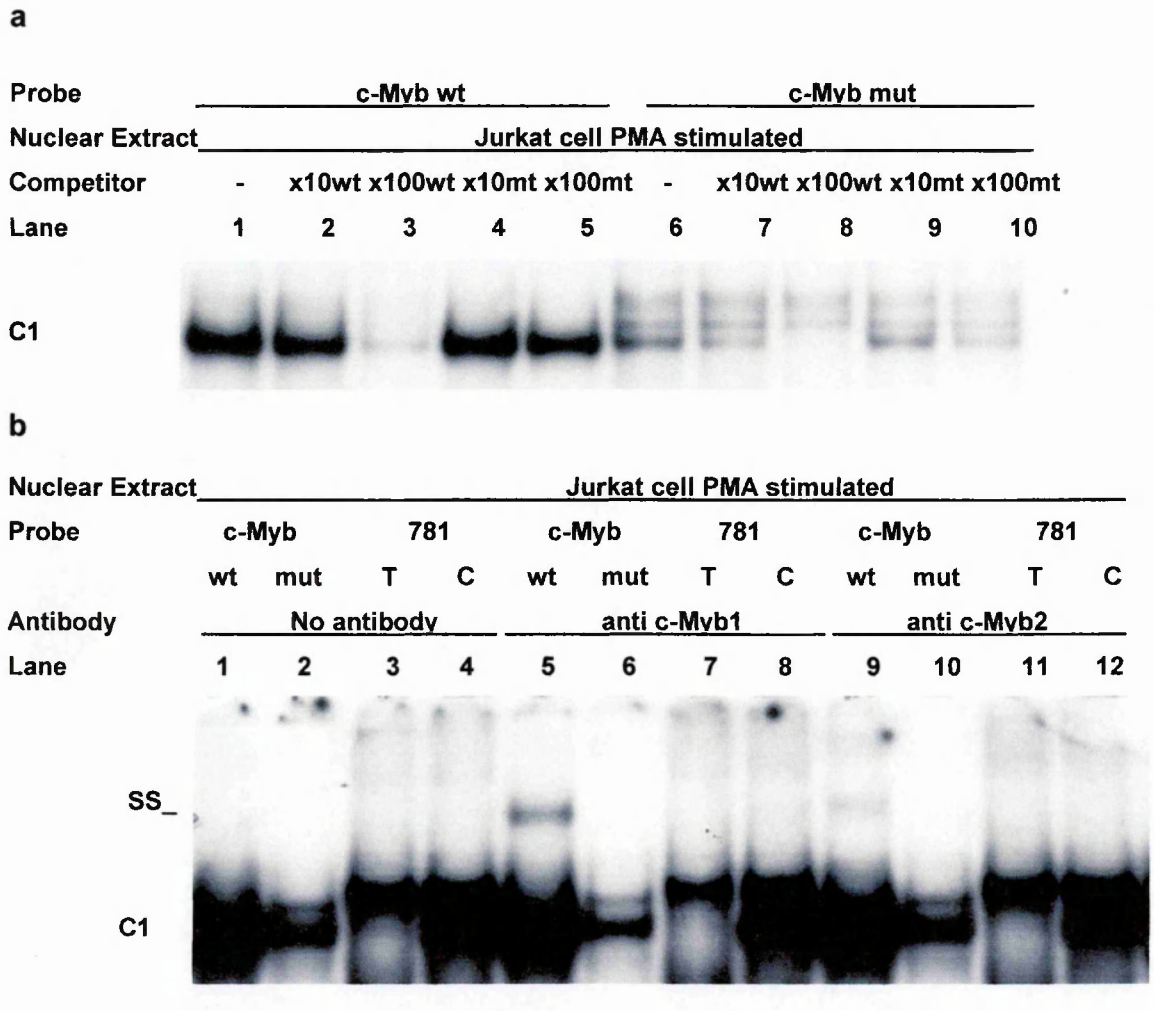


Figure 6.5

c-Myb does not bind in the presence of the +781T SNP in Jurkat cell nuclear extract

Figure 6.5 **c-Myb does not bind in the presence of the +781T SNP in Jurkat cell nuclear extract.** (**a**) Probes were used corresponding to the c-Myb consensus sequence (wildtype, wt) (*lanes 1* to *5*) and the mutated c-Myb consensus sequence (mutant, mt) (*lanes 6* to *10*) in EMSA with nuclear extracts from Jurkat cells harvested after stimulation with PMA. Unlabelled competitor oligoduplex probes were added in molar excess as shown in *lanes 2* to *5* and in *lanes 7* to *10* (x10wt and x100wt refer to x10 and x100 molar excess of unlabelled c-Myb wildtype probe respectively; x10mt and x100mt refer to x10 and x100 molar excess of unlabelled c-Myb mutated probe respectively). Complex *1* (**C1**) bound in a specific manner to the c-Myb wildtype probe as it was competed only by 100 molar excess of c-Myb wildtype competitor probe (*lane 5*). (**b**) c-Myb wt (*lanes 1, 5* and *9*), c-Myb mt probes (*lanes 2, 6* and *10*), +781T probes (*lanes 3, 7* and *11*) and +781C probes (*lanes 3, 7* and *11*) were incubated in the absence (*lanes 1* to *4*) and presence of anti c-Myb antibodies (anti c-Myb1: clone 1-1: *lanes 5* and *8*; c-Myb2: M-19: *lanes 9* to *12*). Complex *1* (**C1**) is present after incubation with the c-Myb wt probe (*lane 1*). Supershifted bands (indicated by **SS_**) formed with both anti c-Myb antibodies are seen only in the presence of the c-Myb wt probe (*lanes 5* and *9*).

6.2.4

C/EBP β is part of the complex binding in the presence of the +781T SNP

Inspection of the sequence surrounding the +781 SNP showed that a C/EBP β consensus binding site (5'-RTTGCCGYAAY-3', where R= A or G and Y= C or T) (Osada et al 1996) was present on the forward strand of the +781T allele (GTTGTTCAAT) but was disrupted on the +781C allele (GTCGTTCAAT). I tested whether C/EBP β was present in the nuclear extracts of A549 cells. As a control I looked for C/EBP β in Jurkat cells as well as in COS cells transfected with a vector expressing C/EBP β (Figure 6.6). C/EBP β was found in both A549 and Jurkat nuclear extracts.

I tested whether C/EBP β was a component of the allele-specific constitutively binding complex *I* seen on the +781T probe in A549 cells by using specific antibodies to C/EBP β. I observed a partial reduction in complex *I* band intensity and a shifted band (Figure 6.7a, *lane 2*), whereas incubation of anti-C/EBP β antibody with the +781C probe had no effect (*lane 4*). By contrast, when this experiment was repeated in Jurkat cells (Figure 6.7b) there was no supershifted band present after incubation of the +781T probe with anti C/EBP β antibodies (lane 2) despite the presence of C/EBP β in the nuclear extracts. As there was only a partial reduction in the complex *I* band binding to the +781T probe in A549 nuclear extract (Figure 6.7a) and no evidence of C/EBP β binding in Jurkat nuclear extract we sought to confirm C/EBP β specific binding to the +781T probe.

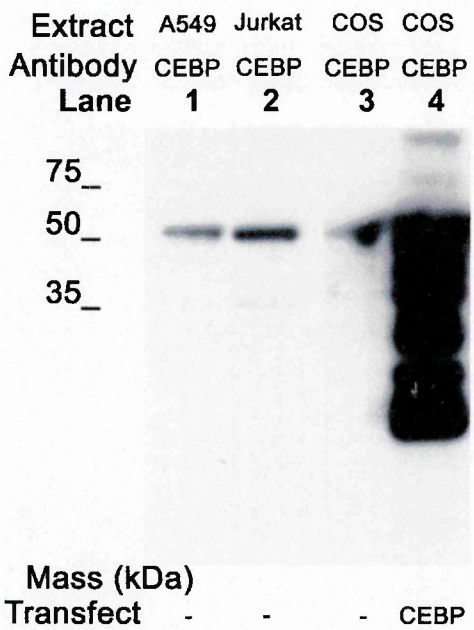
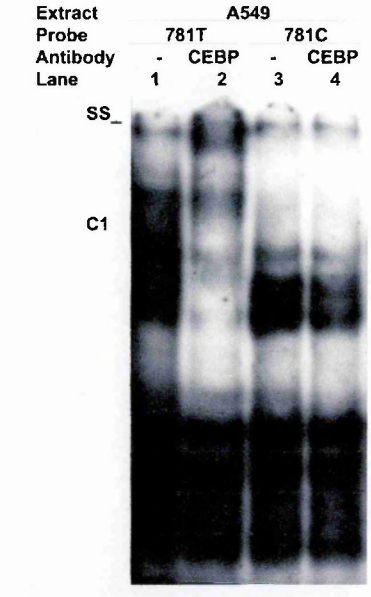


Figure 6.6. **C/EBP β is present in nuclear extracts from A549 and Jurkat cells.** A 50 kDa band was detected on assay by Western blot for C/EBP β in A549 (*lane 1*) and Jurkat (*lane 2*) cell nuclear extract. C/EBP β was also detected in untransfected COS-7 cell nuclear extracts (*lane 3*) and strongly detected in transfected COS-7 cell nuclear extract (*lane 4*). The molecular mass (in kDa), determined by co-electrophoresed standards, is indicated on the left.



b

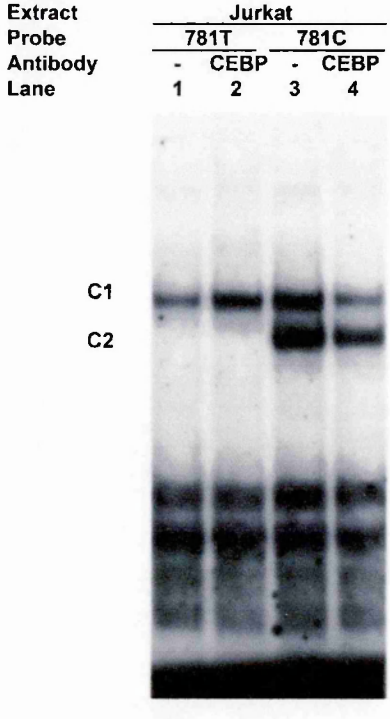


Figure 6.7 C/EBP β binds in the presence of the +781T allele but not the +781C allele in A549 but not Jurkat nuclear extracts

Figure 6.7 **C/EBP β binds in the presence of the +781T allele but not the +781C allele in A549 but not Jurkat nuclear extracts** (**a**) EMSA showing nuclear extract from A549 cells after incubation with probe in the absence (*lanes 1* and *3*) or presence of antibodies (*lanes 2* and *4*) to C/EBP β. Complex *I* (**C1**) is present after incubation with the +781T probe (*lane 1*). A supershifted band (indicated by **SS_**) formed with the anti-C/EBP β antibody is seen only in the presence of the +781T probe (*lane 2*) (**b**) EMSA showing nuclear extract from Jurkat cells after incubation with probe in the absence (*lanes 1* and *3*) or presence of antibodies (*lanes 2* and *4*) to C/EBP β. Complex *I* (**C1**) is present after incubation with the +781T probe (*lane 1* and *2*) whilst Complex II (**C2**) is present after incubation with the +781C probe (*lane 3* and *4*). No supershifted bands (indicated by **SS_**) formed with the anti-C/EBP β antibody.

I performed additional experiments using whole extracts from COS-7 cells previously transfected with a vector expressing C/EBP β (Figure 6.8, *lanes 2, 3, 5, 6*) or an empty vector (Figure 6.8, *lanes 1, 4*). A complex binding to the +781T probe was seen only in the presence of expressed C/EBP β (Figure 6.8, *lane 2*) and this was supershifted by specific antibodies to C/EBP β (*lane 3*). By contrast, little binding of C/EBP β was seen in the presence of the +781C probe when incubated with the same whole COS-7 cell extract containing C/EBP β (*lane 5*). Incubation of nuclear extracts with non-specific antibodies (anti-p65 NF-κB and anti-Oct-1) resulted in no supershifted bands (data not shown).

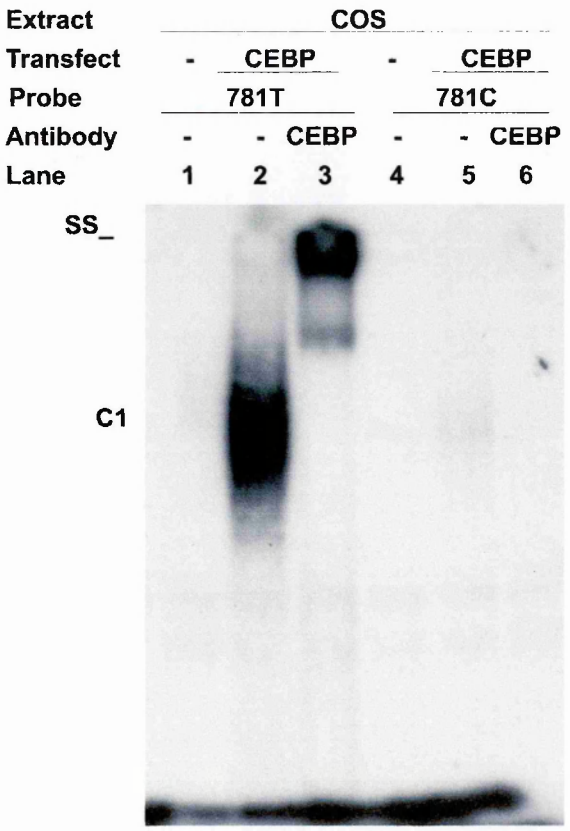


Figure 6.8 Preferential binding of C/EBP β to the +781T allele but not the +781C allele confirmed. Whole cell extract from COS-7 cells was incubated with probe in the absence (*lanes 1, 2, 4* and *5*) or presence of antibodies to C/EBP β (*lanes 3 and 6*). Cells were transfected with a plasmid construct expressing C/EBP β (*lanes 2, 3, 5* and *6*) or an empty vector (*lanes 1* and *4*). Complex I (**C1**) is seen only in the presence of both the +781T probe and C/EBP β protein expression (*lane 2*) and is supershifted significantly (indicated by **SS_**) when incubated with anti-C/EBP β antibody (*lane 3*).

Discussion

6.3.1

Through a systematic examination of the DNA-protein interactions associated with the five SNPs in the *IL8* gene I have shown that there is specific differential binding only between the +781T and +781C alleles. I have demonstrated that C/EBP β was a part of the complex *I* which bound in a constitutive manner specifically in the presence of the +781T allele. These data suggest that C/EBP β may bind to haplotype 2 which is associated with susceptibility to RSV-induced bronchiolitis but not to haplotypes 1 and 3 (Figure 1.6) which have no disease association. Moreover, C/EBP β did not bind to the +781T probe in Human T lymphoblastic leukemia Jurkat cell nuclear extract, a finding which is compatible with the observation that there was no haplotype-specific expression of *IL8* in primary lymphocytes.

In screening the five polymorphisms in the *IL8* gene for DNA-protein interactions we acknowledge that this is strictly a test of differential binding between alleles and does not test for function at an SNP. Furthermore the allele-specific nature of that binding occurs in the absence of the chromatin and the normal architecture of the nucleus. Therefore, more data is required on the specificity and functionality of C/EBP β binding to the +781T polymorphism *in vivo*. Moreover, further experiments would be required to elucidate why despite the presence of C/EBP β in both A549 and Jurkat cell nuclear extract *in vitro* binding of C/EBP β only occurred in the former.

6.3.2

Predictions of c-Myb binding to the +781T allele

In order to identify candidate transcription factors that would exhibit allele-specific binding to the +781 polymorphism we used predictions from the MATINSPECTOR/TRANSFAC program. The suggestion that the c-Myb factor could bind to the +781T allele alone was biologically plausible. The c-Myb transcription factor is a sequence specific DNA-binding protein that plays an essential role in the development of the haematopoietic system (Graf 1992) being expressed at high levels in the immature haematopoietic cells and down regulated at terminal differentiation. However, c-Myb also functions in non-haematopoietic cells as it is present during mouse development in the neural retina and airway epithelium and remains at high expression in adult tissues associated with rapid proliferation such as the hair follicles and gastrointestinal crypts (Sitzmann et al. 1995; Ess et al. 1999). The involvement of c-Myb in developing airway cells was of particular interest to our investigation of *IL8* regulation in the airways of infants. In addition to its role in differentiation c-Myb, in association with other regulatory factors such as the CAAT-enhancer protein family (C/EBP), regulates genes relevant to the immune response such as lysozyme (Ness et al. 1993), myeloperoxidase (Bristos-Bray and Friedman, 1997) neutrophil elastase (Oelgeschlager et al. 1996) and cytokines (Kowenz-Leutz et al. 1997).

Contrary to the predictions of the MATINSPECTOR/TRANSFAC program (Wingender et al. 2000) neither c-Myb nor the ubiquitously expressed B-Myb, with which it shares the same DBD, bound to the +781T probe. I acknowledge that the MATINSPECTOR/TRANSFAC profile predicts the likelihood that the sequence in question is a binding site rather than the binding affinity at that site. Quantitative prediction

of DNA-protein interactions using principal coordinates analysis systematically fitted to experimental data has been shown to have greater accuracy than the profile based models (Udalova et al. 2002) and could be of use in assessing other transcription binding sites at or close to SNPs of interest.

6.3.3

C/EBP β binding to the +781T allele

It was the close association of c-Myb to C/EBP β which lead me to inspect the sequence close to the +781 polymorphism more closely for potential C/EBP β binding sites. C/EBP β is an important candidate transcription factor in influencing RSV-induced bronchiolitis pathogenesis as it has an important general role in the regulation of inflammation (Poli 1998) and is strongly induced by inflammatory stimulants such as TNF (Akira et al 1997) and RSV (Jamaluddin et al 1996). Detailed studies of *IL8* expression have shown that AP-1, IFN regulatory factor (IRF)-1 and C/EBP β are all necessary for RSV-induced induction of the *IL8* promoter (Casola et al 2000, Mastronarde et al 1996). Moreover NF-κB and C/EBP β binding at the promoter synergistically activate *IL8* in response to TNF (Matsusaka et al 1993). Intriguingly, I found that C/EBP β binds to the haplotype that has significantly increased transcription. To our knowledge recruitment of C/EBP β has not previously been modulated by an intronic polymorphism. It is possible that C/EBP β could influence transcription from this location either by interacting with NF-κB through long-range interactions with the promoter as has been shown for c-Myb (Tahirov et al 2002) or through interaction with components of the RNA transcript-

associated heterogeneous nuclear ribonucleoprotein (hnRNP) complexes such as hnRNP K (Miau et al 1998).

6.3.4

Concluding Points

- On systematic screening for allele-specific binding of nuclear factors to SNPs within the *IL8* gene significant differential binding was observed only at the +781 polymorphism.
- This allele-specific binding was seen at the +781 polymorphism with both RSV stimulated A549 cell nuclear extract as well as TNF stimulated nuclear extract.
- c-Myb was not demonstrated to be a part of the complexes binding in the presence of the +781T SNP.
- C/EBP β was found in both A549 and Jurkat nuclear extracts.
- C/EBP β binds in the presence of the +781T allele but not the +781C allele in A549 nuclear extracts and when over-expressed in COS whole cell extracts.
- Notably C/EBP β is not part of the complex binding in the presence of the +781T SNP in Jurkat nuclear extracts.
- These data suggest that C/EBP β may bind to haplotype 2 which is associated with susceptibility to RSV-induced bronchiolitis but not to haplotypes 1 and 3 which have no disease association.
- C/EBP β did not bind to the +781T probe in Human T lymphoblastic leukemia Jurkat cell nuclear extract, a finding which is compatible with the observation that there was no haplotype-specific expression of *IL8* in primary lymphocytes.

CHAPTER SEVEN

DEMONSTRATION OF REGULATORY POLYMORPHISMS MODULATING HAPLOTYPE-SPECIFIC EXPRESSION USING HaploChIP

7.1.1

Introduction

In the preceding chapters I have established that *IL8* exhibits haplotype-specific expression which suggests a functional explanation for the association between haplotype 2 and RSV-induced bronchiolitis disease severity. Moreover, within haplotype 2 there is allele-specific binding of a transcription factor, C/EBP β known for its role in immune regulation. However the extent to which functional effects at an individual SNP within the *IL8* gene can be established is limited by a number of technical factors. Firstly, the European population structure of the *IL8* gene consists of a mirror haplotype. Examination of this structure gives little indication of which SNPs are likely to be functional *in vivo*. Secondly, only one of the SNPs is present within the transcript. This prevents verification of *in vivo* haplotype-specific expression with a second marker. Thirdly, five of the six SNPs in the haplotype are located after the translational start site of the gene. This means that *in vitro* transfection experiments would use constructs that tested SNP function outside the normal sequence of regulatory elements.

The central aim of this thesis was to explore the extent to which SNPs, alone or in combination, influence the transcriptional regulation of genes that play a critical role in

the immune response. Given the limitations to the *IL8* and RSV-induced bronchiolitis model I sought to progress the central objective of the thesis using a different model system. I wished to demonstrate the principle of SNP mediated gene regulation *in vivo* using a novel technology developed in our laboratory, namely haplotype-specific chromatin immunoprecipitation (haploChIP) (Knight et al 2003). This technique was applied to the tumor necrosis factor/lymphotoxin-α (*TNF/LTA*) locus as there was data demonstrating haplotypic diversity in this region. It was our hope that once the utility of haploChIP was demonstrated at the TNF/LTA locus we could apply this technique to the problem of understanding what effect the non-coding SNPs may have on modulating *IL8* transcription.

The haploChIP technique has proved to be a sensitive method for detecting SNPs that affect gene regulation *in vivo*. Moreover the haploChIP method can do so even when there is no suitable genetic marker within the transcript as haploChIP uses the binding and activation of transcriptional proteins as a surrogate maker of transcription. For instance, for synthesis of the transcript to occur RNA polymerase II (Pol II) must be phosphorylated on specific serine residues in the C-terminal domain (CTD) after which the enzyme is released from the initiation complex. Within the CTD, phosphorylation of the Ser5 marks the transition from mRNA initiation to elongation (Komarnitsky et al 2000). This is in constrast to phosphorylation at Ser2 which occurs at a later stage in elongation (Cho et al 2001). CTD phosphorylation of Ser5 denotes also the beginning of RNA capping (Schroeder et al 2000). The importance of phosphorylation within the CTD has been verified by experiments which show that that the amount of phosphorylated Pol II detected within chromatin is correlated with the transcription at that gene (Weeks et al 1993).

Critical to the flexibility of haploChIP is the ability of this technique to demonstrate allele-specific expression in genes where there are no transcribed markers. As shown in Figure 7.1, the haploChIP technique starts by crosslinking the activated Pol II protein to the DNA through the use of formaldehyde. Then through sonication the chromosomes are sheared to DNA fragments with a mean length of less than 3 kilobases. The DNA fragments attached to Pol II are then retrieved through immunoprecipitation using antibodies specific for the Ser5 and Ser2 CTD forms of Pol II. These immunoprecipitated DNA fragments are then amplified by PCR and distinguished by primer extension (PE). Then, the relative abundance of the two alleles, which were bound to Pol II, are determined by MS. The size of the DNA fragments determines the distance at which the marker can be detected through the immunoprecipitation of activated Pol II. The technique is likely to be more sensitive where the genetic marker and the cross-linked protein are close to each other.

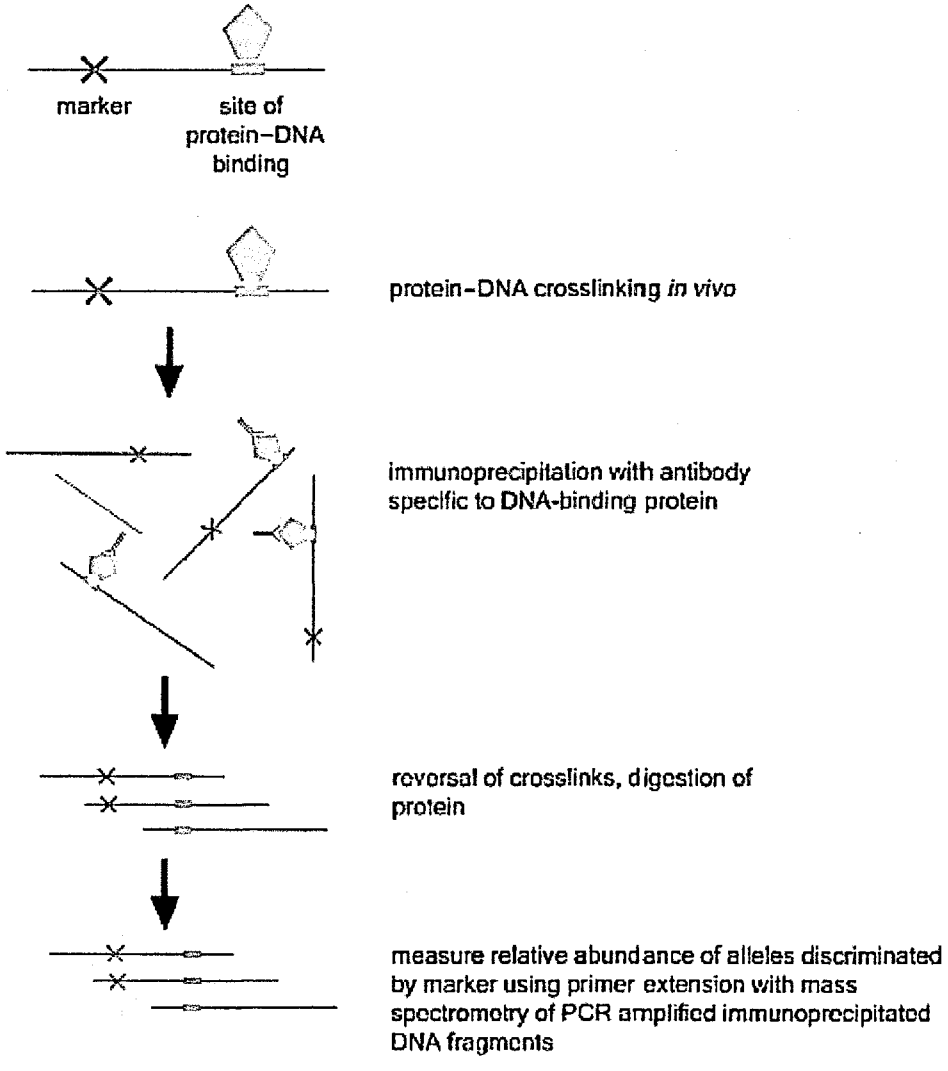
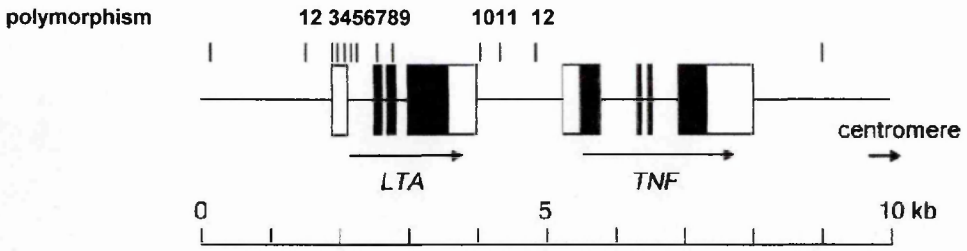


Figure 7.1. **Schematic diagram demonstrating the quantitation of differential protein-DNA binding *in vivo* by HaploChIP.** DNA binding proteins are cross linked to the DNA with formaldehyde after which the DNA is sheared through sonication. Fragments bearing cross-linked proteins of interest are retrieved by immunoprecipitation. The bonds between the proteins and antibodies are reversed after which the crosslinks between proteins and the DNA are broken. The relative abundance of alleles can then be measured by PE/MS. Diagram reproduced from Knight and co-workers, Nature Genetics 2003 and used with permission from the author.

The HaploChIP approach was used to show differences in the amount of Pol II bound to different alleles of *LTA* encoding lymphotoxin which were closely correlated with allele-specific expression (Knight et al 2003). In this instance an allele-specific difference was observed by quantitation of *LTA*+252 (through haploChIP) and of the marker *LTA*+723 found in the *LTA* transcript. The haplotype carrying the *LTA*+252G and *LTA*+723A alleles could be functional as it was associated with higher *LTA* expression. However, closer inspection of the data reveals that the *LTA*+252G allele may not in themselves be functional. Instead, the allele-specific function may derive from other differences between the most common haplotype that did not contain the *LTA*+252G allele and the other haplotype that did (Figure 7.2). Further detailed analysis of the *TNF/LTA* locus identified the *LTA*+80A allele as being associated with reduced LTA protein production in human EBV immortalised B cells (Knight et al 2004). In addition chromatin immunoprecipitation confirmed that Activated B cell Factor-1 (ABF-1) bound only in the presence of the low producing *LTA*+80A allele *in vivo* (Knight et al 2004). I sought to confirm this finding, in collaboration with Julian Knight, by conducting a detailed analysis of the *TNF/LTA* haplotype clades described (Figure 7.2) using haploChIP. We hoped to demonstrate that there was reduced transcription at the *LTA*+80A allele in accordance with the suppression of protein production that has previously been reported. If demonstrated this observation would be consistent with the conclusion that SNPs acting within a haplotype could modulate gene regulation *in vivo* at the level of transcription and that these haplotype specific differences in transcript cause significant variation in the protein production.



	Polymorphism			Haplotype						
	SNP	Minor Allele Frequency	Sequence		Clade A		Clade B		Clade C		
				Haplotype Frequency	A1 0.096	A2 0.173	B1 0.423	B2 0.038	C1 0.154	C2 0.077	C3 0.038
1	*LTA*-294	0.154	AAC^{C}/$_{T}$GCC		C	C	C	C	T	C	C
2	*LTA*-293	0.423	ACC^{G}/$_{A}$CCT		G	G	A	G	G	G	G
3	*LTA*+10	0.269	TCC^{G}/$_{A}$CAC		A	A	G	G	G	G	G
4	*LTA*+80	0.461	CCC^{C}/$_{A}$GCA		C	C	A	A	C	C	C
5	*LTA*+252	0.269	ATG^{A}/$_{G}$TTC		G	G	A	A	A	A	A
6	*LTA*+368	0.461	GCT^{G}/$_{C}$TCT		G	G	C	C	G	G	G
7	*LTA*+438	0.269	CTG(TCn)		9	9	9	9	10	10	10
8	*LTA*+495	0.269	GTG^{T}/$_{C}$GTG		T	T	T	T	C	C	C
9	*LTA*+723	0.269	GCA^{C}/$_{A}$CCT		A	A	C	C	C	C	C
10	*TNF*-1031	0.231	AGA^{T}/$_{C}$GAA		T	T	T	T	C	C	T
11	*TNF*-863	0.154	CCC^{C}/$_{A}$CTT		C	C	C	C	A	C	C
12	*TNF*-308	0.173	ATG^{G}/$_{A}$GGA		G	A	G	G	G	G	G

Figure 7.2. **Haplotype structure at the *TNF/LTA* locus.** Cloning and sequencing of a 10-kb region of the locus spanning *TNF* and *LTA* revealed 14 SNPs. Twelve of the SNPs that occurred at an allele frequency of greater than 0.1 in the 26 lymphocyte lines tested are shown together with their corresponding haplotypes. Minor allele frequencies are shown as well as the haplotype frequencies at the top of the clade columns. Designation of the polymorphisms refers to the nucleotide position relative to the translational start site. SNPs that uniquely define clades are shown as follows (white lettering): Clade A is defined by *LTA*+10A, *LTA*+252G and *LTA*+723A; Clade B is defined by *LTA*+80A and *LTA*+368C; and Clade C *LTA*+438(10), *LTA*+495C and *TNF*-1032C. Adapted published figures (Knight 2003, 2004) used after permission from the author.

Results

7.2.1

Optimisation and verification of the haploChIP method for high through put chromatin production

In order to undertake a systematic investigation of haplotype-specific expression at the *TNF/LTA* locus 23 EBV immortalised B cell lines from the CEPH repository were cultured. These lines were then stimulated with PMA/Ionomycin to initiate *LTA* expression and then cross-linked with formaldehyde at time points of 0, 30, 120 and 360 minutes. To generate sufficient chromatin for the study the standard protocol for haploChIP (Knight et al 2003) had to be scaled up by four fold. This necessitated a number of changes to the standard protocol.

Firstly, the sonication equipment and conditions were altered. The micro-tip used previously was changed to a medium-tip so that efficient sonication would occur in a larger volume of extracted nuclear suspension. In addition, the sonication regime was changed from 30 second bursts of which six were at setting 4 and six at setting 6. Here 40 cycles (4 second burst with 8 seconds pause) at an amplitude of 20% followed by 40 cycles (4 second burst with 8 seconds pause) at an amplitude of 30% were used. I sought to ensure that this modified sonification strategy provided fragments of the expected average length of less than 3 kilobases. As shown in Figure 7.3, sonication under these conditions produced fragments ranging from under 400 base pairs to over 10,000 base pairs. However, the most significant band was seen at about 1.5 kilobases.

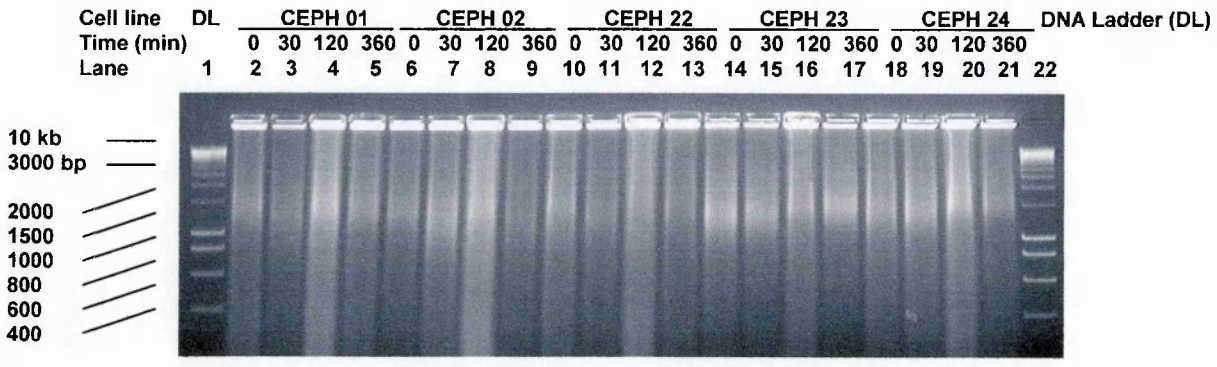
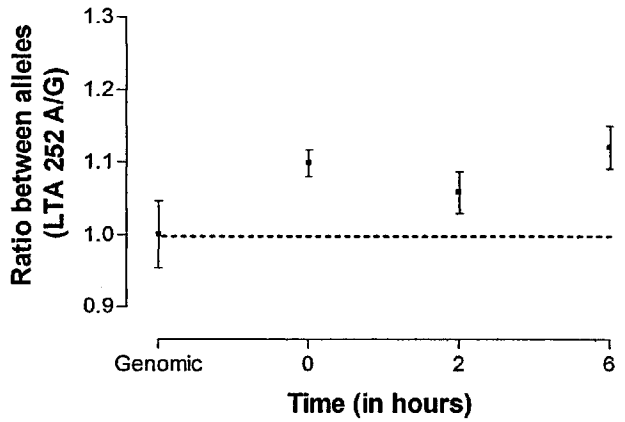


Figure 7.3. **Demonstration of DNA fragments size after sonication of extracted nuclear suspension.** A sample of DNA fragments obtained from nuclear extraction of selected CEPH EBV immortalized B cells is shown. Samples were diluted with N-laurosarcosine (Sigma) to 0.5% and run on a 1.4% TBE agarose gel. Electrophoresis of standard size markers of DNA (DNA ladder, DL) is displayed in lanes 1 and 22. In all cases fragments vary in size between 400 base pairs (bp) and over 10 kilobases (kb) with the main band occurring at about 1.5 kilobases.

Secondly, the requirement for a Caesium Chloride (CsCl) column for final purification of the chromatin fragments was likely to be a significant rate limiting step to the processing of larger volumes of chromatin as ultra-centrifugation over twenty hours was required. In addition limited volumes of material could be placed on the CsCl column. Chromatin was prepared for immunoprecipitation with and without CsCl purification. The resulting DNA was tested for allele-specific expression to ensure that the previously published observation of an effect associated with the *LTA*+252 SNP was reproduced with or without CsCl column purification. As shown in Figure 7.4, there appears to be significant allele-specific expression associated with the *LTA*+252 polymorphism when using DNA fragments that have been purified both with and without CsCl column purification. Within this cell line, however, the allelic ratio shows that greater *LTA* transcription is associated with the *LTA*+252A allele and not the *LTA*+252G allele. Based on this experiment there appears to be no difference between DNA fragments prepared by CsCl column purification and those which are not. Therefore the CsCl column purification was dropped from the protocol. However, the association of greater *LTA* transcription with the *LTA*+252A allele as compared with *LTA*+252G in the context of haplotypes C and A respectively, is a novel observation. This observation does not contradict recent findings suggesting that the *LTA*+80A allele found in haplotype B has a suppressor effect on LTA protein production (Knight et al 2004).

a



b

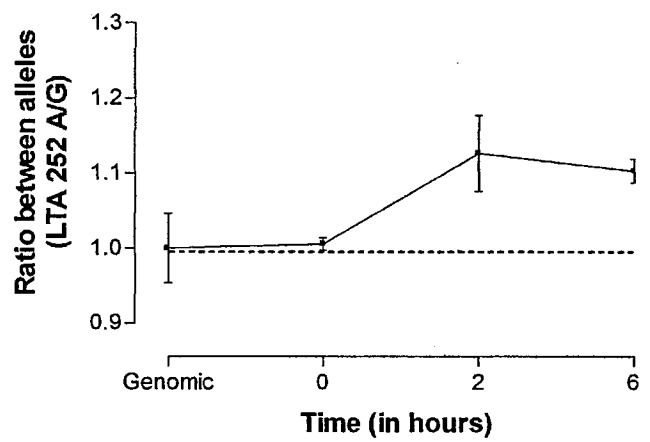


Figure 7.4. **Allele specific loading of phosphorylated Pol II *in vivo* at *LTA* after chromatin preparation with and without CsCl purification.** The ratio of phosphorylated Ser 5 Pol II loading between alleles A and G at the *LTA*+252 SNP is shown for a single B cell line, heterozygote for haplotypes A and C, after stimulation with PMA/ionomycin over time. DNA fragments were purified with (**a**) and without (**b**) CsCl column purification. In each case the mean ± the standard error of the mean is shown for four independent PCR amplification reactions each of which was spotted four times on a chip for PE/MS. Genome represents the *LTA*+252 allelic ratio for genomic DNA.

7.2.2

LTA+80, LTA+367 and LTA+723 are independently associated with allele-specific polymerase loading

It has been previously shown (Knight et al 2003, Knight et al 2004) that twelve polymorphic sites present within the *TNF/LTA* locus, at an allele frequency of greater than 10% could be used to define three major haplotype clades across the *TNF/LTA* locus (Figure 7.2). For the work presented in this chapter six SNPs were used, namely *LTA*+10, *LTA*+80, *LTA*+252, *LTA*+368, *LTA*+723 and *TNF*-1031, as markers for the haplotype clades. In order to investigate the relationship between genetic variation at the *LTA/TNF* locus and Pol II loading the 23 B cell lines previously cultured, stimulated and cross-linked were typed to determine their haplotype-clade. Of the 23 B cell lines 16 were shown to be heterozygous for two of the three clades under study. Haplotype-specific Pol II loading at 0, 30, 120 and 360 minutes after PMA/ionomycin stimulation was then assayed using haploChIP. In conducting these experiments it was important to control for experimental variation between SNP assays across cell lines and for systematic allelic difference in PE/MS signal intensity. To do this we paired genomic DNA and haploChIP DNA from the same cell line for assay of each SNP by PE/MS. The genomic DNA ratio was used to prove a correction factor for systematic differences in signal intensity for each assay. This correction factor was applied equally to both genomic DNA and haploChIP DNA. To account for variation between experiments the haploChIP DNA ratio at each time point was divided by the mean genomic DNA ratio for each assay within that cell line.

In the first instance univariate linear regression was used to investigate whether any the six marker SNPs were significantly associated with haplotype-specific loading of Pol II. Of the six SNPs assayed only three SNPs, *LTA*+80 ($p<0.0001$), *LTA*+367 ($p=<0.0001$)

and *LTA*+723 (p=<0.0001) were significantly associated with allele-specific expression. However, the high linkage disequilibrium across the *LTA* locus means that an SNP on univariate analysis could appear to be functional when in fact it is acting as a marker for an adjacent SNP to which it is closely linked. To test whether these three SNPs had independent effects on Pol II loading *LTA*+10, *LTA*+80 and *LTA*+723 were tested together in a multivariate regression model. As shown in Table 7.1, all three appeared to be independently associated with allele-specific expression of *LTA*.

SNP	Coefficient	Standard Error	95% Confidence interval	p value
LTA+80	-0.102	0.029	-0.159, -0.044	0.001
LTA+367	0.181	0.038	0.107, 0.255	<0.0001
LTA+723	-0.182	0.034	-0.249, -0.115	<0.0001

Table 7.1. **Multivariate regression model of haploChIP DNA ratio against SNP.** Linear regression of haploChIP (dependent variable) against SNPs in the *LTA* locus which occur at an allele frequency of over 10% and which were independently associated with allele-specific polymerase loading.

7.2.3

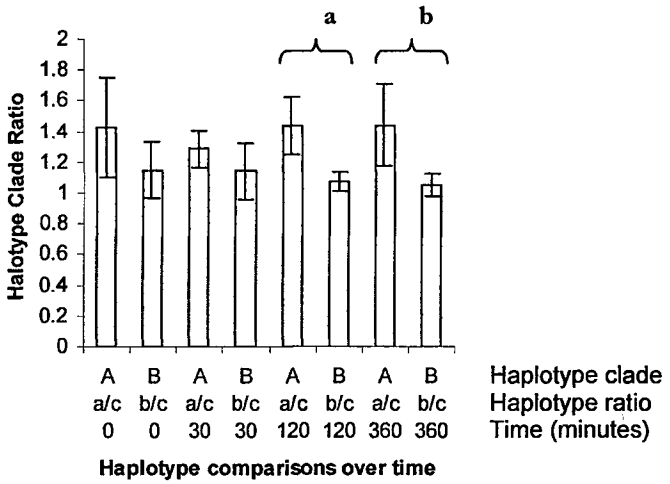
Polymorphisms LTA+80 and LTA+367 are associated with time dependent reduced haplotype-specific polymerase loading

We sought to determine which of the alleles on the *LTA*+80, *LTA*+367 and *LTA*+723 polymorphisms were associated with functional effects. As shown in Figure 7.2 *LTA*+80A and *LTA*+368C define clade B whilst *LTA*+723A is found in clade A and *TNF*-1031C is found in clade C. The haploChIP experiments generated data in the form of haplotype ratios. In order to understand the relationship of haplotype clades A, B and C to each other I compared haplotype ratios that shared a common denominator. Therefore, a comparison of A to B could occur by regression of A/C with B/C. Similarly, A to C can be compared with regression analysis of A/B against C/B.

As shown in Figure 7.5a comparison of haplotype clade B, containing alleles *LTA*+80A, *LTA*+368C and *LTA*+723C, with haplotype clade A (*LTA*+80C, *LTA*+368G and *LTA*+723A) shows that there appears to be significantly reduced polymerase loading associated with clade B at 120 minutes after induction (A/C 1.44, B/C 1.07; p=<0.0001) and at 360 minutes after induction (A/C 1.44, B/C 1.05; p=0.005). This is consistent with previous data that demonstrated a repressor effect associated with *LTA*+80A. By contrast, as shown in Figure 7.5b when clade A (containing *LTA*+723A) was compared with clade C (containing *LTA*+723C) there was no significant differences in polymerase loading at either 120 minutes (A/B 1.05, C/B 0.97; p=0.067) or 360 minutes (A/B 0.99, C/B 1.00; p=0.757). The absence of a difference in polymerase loading between clades A and C confirms the results of the univariate and multivariate analyses which suggested that polymorphisms *LTA*+10, *LTA*+252 and *TNF*-1031 may not have functional genetic properties. Furthermore it implies that the allele *LTA*+723C present in clades C

and B may not have functional properties when separated from the alleles *LTA*+80A and *LTA*+368C found on clade B.

a



b

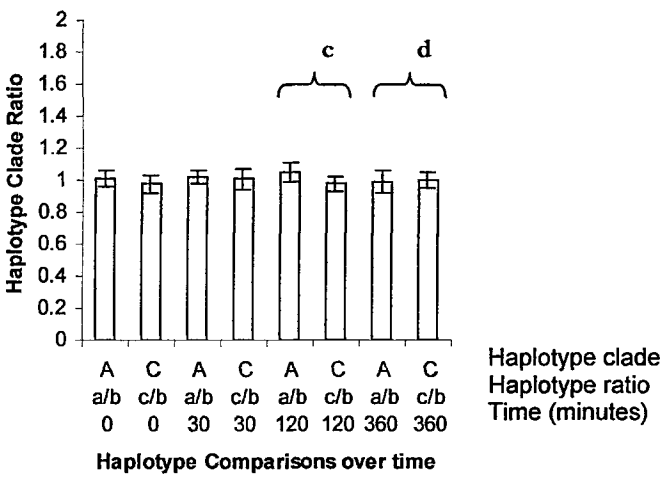


Figure 7.5 Comparison over time of haplotype-specific loading of phosphorylated polymerase II.

Figure 7.5 Comparison over time of haplotype-specific loading of phosphorylated polymerase II. (**a**) Over time haplotype-specific ratios of polymerase loading were measured in cell lines heterozygous for haplotypes A and C and compared to cells heterozygous for haplotypes B and C. Significantly greater polymerase loading was associated with haplotype A when compared to B at 120 ([a]A/C 1.44, 95% CI 1.63, 1.25; B/C 1.07 95% CI 1.00, 1.14; p=<0.0001) and 360 minutes ([b]A/C 1.44, 95% CI 1.71, 1.18; B/C 1.05 95% CI 0.98, 1.13; p=0.005). (**b**) Over time haplotype-specific ratios of polymerase loading were measured in cell lines heterozygous for haplotypes A and B and compared to cells heterozygous for haplotypes C and B. No significant differences in polymerase loading were seen associated with haplotype A when compared to C at either 120 ([c]A/B 1.05, 95% CI 0.99, 1.11; C/B 0.97 95% CI 0.93, 1.02; p=0.067) or 360 minutes ([d]A/B 0.99, 95% CI 0.92, 1.06; C/B 1.00 95% CI 0.95, 1.05; p=0.757). Data is represented as means with 95% confidence intervals.

Discussion

7.3.1

Limitations of the IL8 locus as a model of in vivo haplotype-specific transcription

This thesis aims to test the hypothesis that natural variation in the genome, in the form of single nucleotide polymorphisms, influenced the regulation of genes critical to the immune response. A mechanism for this would be by influencing gene regulation at the level of transcription by changing the affinity with which transcription factors bind to that site. I reasoned that this small but significant alteration in gene expression produced differences in the protein production that would alter the individual's susceptibility to common disease states. Exploration of *IL8* gene regulation in the context of RSV-induced bronchiolitis disease severity has demonstrated some but not all of these principles. RSV-induced bronchiolitis disease severity is linked with increased lung inflammation. I have shown that a haplotype of *IL8* associated with host disease susceptibility has significantly greater expression in primary respiratory cells *in vivo* and that this haplotype is associated with allele-specific binding of C/EBP β *in vitro.* However, I do not have data on how IL-8 protein production in respiratory cells relates to the two most common haplotypes in the European population. Moreover, I have not been able to resolve with certainty which of the SNPs within the *IL8* haplotype, acting alone or in concert, have functional properties *in vivo.* I reasoned that the HaploChIP technique may be able to resolve the extent to which non-coding polymorphisms modulate *IL8* transcription. However, prior to applying HaploChIP to *IL8* I sought to optimise the method at the *LTA/TNF* locus. Having achieved this aim I proceeded to explore the extent to which non-coding polymorphisms affect the previously demonstrated haplotype-specific expression seen at the *LTA* gene.

7.3.2

HaploChIP at the LTA locus as model of in vivo haplotype-specific transcription

However, the study of the *TNF/LTA* locus as a model system has served to demonstrate the central hypotheses of the thesis more fully. The *LTA*+252 SNP has been associated with both susceptibility to infection (Struber et al. 1996, Majetschak et al. 1999) and inflammation (Yamaguchi et al. 2001, Badenhoop et al. 1992) as well as other conditions such as asthma (Albuquerque et al. 1998, Moffatt et al 1997) and myocardial infarction (Ozaki et al. 2002). Studies of transcript and polymerase loading suggest that there appears to be increased allele-specific transcription at the *LTA* locus associated with the haplotypes carrying the *LTA*+252G allele (Knight et al. 2003). Further study of the *LTA* locus refined the functional location at the *LTA* locus to the *LTA*+368C and *LTA*+80A alleles and showed that these SNPs appeared to determine reduced *LTA* protein production (Knight et al. 2004). Moreover, it was demonstrated that the transcriptional repressor ABF-1 was recruited to bind in the presence of the *LTA*+80 polymorphism both *in vitro* and *in vivo*. Finally, the recruitment of ABF-1 has been shown to suppress reporter gene expression *in vitro* (Knight et al. 2004). The findings presented here confirm that the *LTA*+80A and *LTA*+368C polymorphisms are functional determinants of haplotype-specific LTA protein production but more critically demonstrate that this haplotype-specific regulation occurs *in vivo* at the level of transcription. The time dependent nature of the reduced polymerase loading associated with the *LTA*+80A and *LTA*+368C polymorphisms is compatible with the recruitment of transcriptional repressors such as ABF-1 to that haplotype. These data serve also to demonstrate that the *LTA*+10, *LTA*+252 and the *TNF*-1031 polymorphisms do not appear to have a functional role.

The role of the *LTA*+723 polymorphism in governing haplotype-specific *LTA* protein production is less certain. Previously, the *LTA*+723 was found to be individually associated with LTA protein production. However, this association did not remain true when the *LTA*+723 polymorphism was analysed within the context of a stepwise linear regression model (Knight et al. 2004). In this study the *LTA*+723 has been associated with allele-specific polymerase loading in both univariate and multivariate models. This suggests that the association with polymerase loading detected at this locus is not merely a function of the fact that the *LTA*+723C allele is closely linked to the *LTA*+80A and *LTA*+368C polymorphisms. However, on testing the relationship of polymerase loading to the haplotype clades the *LTA*+723C allele did not appear to exhibit a functional effect when tested against the *LTA*+723A allele seen in haplotype clade A. This suggests that the *LTA*+723C allele only becomes functional when on the same haplotype as the *LTA*+80A and the *LTA*+368C polymorphisms. It is possible therefore that in addition to influencing the binding of transcriptional factors such as ABF-1 these three polymorphisms act in concert to influence the binding of other factors necessary for the regulation of transcription. These data suggest therefore that the functional unit for modulation of gene regulation by DNA variation should occur at the level of the haplotype as an analysis based on individual SNPs could prove to be misleading.

7.3.3

Comparison of haploChIP as a measure of function at a haplotype as compared with assays of transcript or protein abundance

An important caveat is that functional genetic properties were detected at the *LTA*+723C allele using polymerase loading, which acts as a surrogate for transcriptional activity, when no such function was established between this allele and protein production. It is possible that there may be differences in the sensitivity of these two

methods to detect function at polymorphisms within haplotype clade B. Given that gene regulation governed by polymorphisms appears to occur at the level of transcription it is possible that the functional properties of polymorphisms are more readily detected when using tests of polymerase loading. This is because haploChIP measures a ratio of RNA polymerase loading between alleles which may be a more accurate reading of transcription than the steady state concentration of transcript or protein at any given time point. Moreover, the functional 'signal' attributed to a polymorphism may be 'dampened' by factors influencing transcript abundance and through other regulatory methanisms affecting translation and protein export. This could mean that polymorphisms that contribute to gene regulation only within the context of other functional polymorphisms may be detected more readily through haploChIP than through assays of transcript abundance or through ELISA studies.

7.3.4

Differences in the examination of haplotype-specific expression in an individual as compared to a population

By including all the haplotypes from the diverse population structure of the *TNF/LTA* locus it is possible to begin to make general inferences about the factors governing haplotype-specific gene expression at this locus. A number of other studies have noted that haplotype-specific expression varies between individuals (Bray et al 2003, Lo et al 2003, Yan et al 2002, Cowles et al 2002). In the optimisation of the haploChIP technique I have shown in one B cell line a consistent increase in allele-specific polymerase loading associated with the *LTA*+252A allele in the context of haplotypes A and C (Figure 7.4). Such allele-specific loading was not demonstrated in a B cell line of the same genotype in previously published data (Knight et al 2003). The difference between studies in allele-specific polymerase loading in cells of the same genotype might

be a function of random experimental variation. However, the design of the studies we have conducted at the *TNF/LTA* locus is highly conservative in that random variation at individual polymorphisms is less likely to produce significant effects when considered within the context of the other polymorphisms that define the three haplotype clades in the chromosomes assayed. Alternatively, differences in allele-specific polymerase loading in cells of the same genotype might imply that there are other epigenetic, host *trans*-acting and environmental factors that also influence allele-specific transcription. However, the detection of haplotype-specific influences on gene expression across 32 chromosomes suggests that despite the potential presence of other influences in these experiments gene regulation through polymorphic variation is the predominant factor.

The inclusive approach adopted here for gene expression is in contrast to previously published studies of allele-specific expression in human cells (Yan et al 2002, Bray et al 2003, Lo et al 2003). Yan et al. (2002) examined 13 genes in 96 individuals from the CEPH families. They identified allelic variation in 6 of the 13 genes studied with the ratio of transcripts varying from between 1.3 to 4.3. However, the number of individuals exhibiting such variation ranged from only 3 to 30%. No analysis is made of the significance of allele-specific expression at a given gene across all individuals possessing an informative heterozygous genotype. Similarly, Bray et al. (2003a) screened 19 subjects and found allele-specific differences in seven of the 15 genes expressed in brain tissue. However, in five of the genes assayed allele-specific expression was seen in only one of the 19 subjects. Here again, it is unlikely that allele-specific expression at a gene would be considered significant if it were assessed in the context of the whole study population instead of at the level of the individual. Lo and co-workers used a microarray method called HuSNP oligo array to screen 602 genes from seven individuals for allele-specific

differences in expression. They found that 326 genes (54%) showed preferential expression of one allele in at least one individual. Of the subset of these 23 genes shown in the article 9 genes do not preferentially express the same allele in 80% of the samples assayed. In that same subset 17 out of the 23 genes were not allele-specifically expressed in the majority of the individuals.

The finding that allele-specific expression is consistently observed in a minority of individuals may still yield important information about gene regulation. However, the attraction of genetic epidemiology and functional genetic investigations lies in part in the potential for discovering fundamental mechanisms of genetic regulation around which drug and vaccine development can occur. If functional genetic studies reveal means of genetic regulation that are restricted to a small number of individuals then the scope for applying these findings to the development of a therapeutic intervention is limited.

7.3.5

Limitations and strengths of the haploChIP data demonstrated at the LTA locus

The data presented here in this chapter would be strengthened by further experiments. Firstly, the haplotype-specific effects observed for polymerase loading and protein production should be verified with the addition of data on transcription abundance. Secondly, the role of the *LTA*+368 and *LTA*+723 polymorphisms in influencing expression at haplotype clade B requires further study. The means by which these polymorphisms may interact with transcriptional factors binding in the presence of *LTA*+80, needs to be clarified.

In conclusion, this chapter demonstrates the utility of haploChIP as a means of assaying polymerase loading. It demonstrates that polymerase loading can be a powerful tool in dissecting the functional properties of haplotypes with a diverse population structure. The work demonstrates the principle that natural genetic variation alters the transcriptional regulation of genes critical to the immune response and that this alteration in transcription is consistent with differences in protein expression. Together with the haplotype-specific data on *IL8* shown in Chapter 4, this chapter demonstrates that differences in haplotype-specific expression can be detected across whole study populations with the relevant genotype. It also shows that HaploChIP could be suitably applied in future to the study of gene regulation at the *IL8* gene. This implies that the functional basis of *in vivo* haplotype-specific expression may occur through *cis*-acting regulatory mechanisms that are modulated by polymorphisms in the DNA. Moreover, the haploChIP experiments detailed here suggest that functional polymorphisms may be located not only in the 5' untranslated region of the gene but also within intronic regions.

7.3.6

Concluding Points

- The haploChIP method can be optimised for high through-put chromatin production.
- *LTA*+80, *LTA*+367 and *LTA*+723 are independently associated with allele-specific polymerase loading.
- Haplotype B defined by the polymorphisms *LTA*+80 and *LTA*+367 is associated with time dependent reduced haplotype-specific polymerase loading.
- *LTA*+723C allele on haplotype B may only become functional when on the same haplotype as the *LTA*+80A and the *LTA*+368C polymorphisms.
- These data suggest that the functional unit for polymorphic modulation of gene regulation occurs at the level of the haplotype.
- Given that gene regulation governed by polymorphisms appears to occur at the level of transcription it is possible that the functional properties of polymorphisms are more readily detected when using tests of polymerase loading.
- The demonstration of haplotype-specific expression across whole study populations implies that some of the functional basis of *in vivo* haplotype-specific expression may occur through *cis*-acting regulatory mechanisms that are modulated by polymorphisms in the DNA.

CHAPTER EIGHT

THESIS SUMMARY

8.1

This thesis explores the extent to which natural variation in DNA sequence may influence the transcriptional regulation of genes that play a critical role in the immune response to infection. Using both novel *in vivo* and traditional *in vitro* techniques this thesis seeks to demonstrate that functionally important genetic variation may facilitate our understanding of disease associations. In this thesis I chose to examine Respiratory Syncytial Virus (RSV) infection as it is common and has a robust genetic association with disease severity. The principal determinant of that disease severity is believed to be the host immune response.

An increasing number of studies have shown associations between single nucleotide polymorphisms (SNPs) situated in or close to candidate genes and differences in disease severity. Until recently demonstrating the functional genetic basis for this has been limited by a number of factors. Firstly, studies relying only on *in vitro* assays of protein-DNA interaction and plasmid reporter gene expression take no account of the regulatory nature of the normal chromatin structure of DNA. Secondly, modulation of gene regulation could depend on the interaction of combinations of SNPs on an allele (the haplotype) whereas these same *in vitro* assays tend to analyse SNPs in isolation. In this thesis three methods of *in vivo* allele-specific quantitation of gene expression are used, namely Mutant analysis by PCR and restriction enzyme cleavage (MAPREC), primer extension mass spectrometry (PE/MS) and haplotype-specific chromatin immunoprecipitation (haploChIP). I used these methods

to measure haplotype-specific transcription in living human primary cells. This was achieved by analyzing cells from individuals heterozygous for SNPs that acted as markers to distinguish between alleles. In cells heterozygous for an SNP marker, it was possible to determine the relative differences in expression between the two alleles. I proceeded to try and determine the molecular mechanism for haplotype-specific expression due to specific SNPs using *in vitro* assays of protein-DNA interaction and plasmid reporter gene expression in immortalised cell lines.

IL-8 has been implicated in the pathogenesis of RSV-induced bronchiolitis as it may play a role in the activation and migration of neutrophils that characterize the inflammatory infiltrates seen in the airways of affected individuals. Previously, the Kwiatkowski laboratory has described an association between bronchiolitis disease severity and a common single nucleotide polymorphism 251 base pairs (bp) upstream (-251) upstream of the *IL8* transcription start site. Further studies defined this finding further by demonstrating that a specific *IL8* haplotype comprising six SNPs (-251A/+396G/+781T/+1238delA/+1633T/+2767T, haplotype 2) was associated with RSV-bronchiolitis disease severity. This second study suggested that the –251 SNP may not be functional and that the association was refined by the addition of the +781 SNP.

In this thesis I investigated the functional basis for this association by measuring haplotype-specific transcription *in vivo* through MAPREC and PE/MS in human primary cells. When assayed by PE/MS I found a significant increase in the transcript level derived from the *IL8* haplotype 2 relative to the mirror haplotype 1

(-251T/+396T/+781C/+1238insA/+1633C/+2767A). This difference in haplotype-specific expression was seen in respiratory epithelial cells but not in lymphocytes.

I proceeded to investigate the extent to which an immortalized cell line, heterozygote for haplotypes 1 and 2, was an adequate model for *in vivo* allele-specific expression. Contrary to the result seen in primary respiratory cells the Human H460 pulmonary epithelial cell line (H460), showed increased haplotype-specific expression of haplotype 1. However, on examination of the H460 karyotype with fluorescent *in situ* hybridization it become clear that there were gross irregularities in both the chromosome 4 number and in the *IL8* copy number in the H460 genome. The results of *in vivo* haplotype-specific expression studies in this cell line were therefore difficult to interpret accurately.

The -251 polymorphism is the only SNP in the haplotype to occur prior to the translational start site of *IL8*. I used *in vitro* reporter gene experiments containing the –251 polymorphism, in the context of the +1409 nucleotides prior to the transcriptional start site, to examine this SNP for allele-specific differences in reporter gene expression in two cell lines. A significant increase in transcription was associated with the -251T polymorphism in H460 but not in A549 cell lines. Furthermore differences in protein-DNA binding were seen at the -251 polymorphism between the H460 and A549 cell lines. The results of the transcription experiments in H460 cells were consistent with the haplotype-specific data derived from the same cells. However, as neither the haplotype-specific data nor the transfection experiments in H460 cells accurately reflected what was seen in the primary respiratory cells this cell line was not used further.

For the remaining SNPs that were located beyond the translational start site of *IL8* I tested the hypothesis that the observed differences in haplotype-specific expression seen in primary respiratory cells could be due to differential binding of transcription factors at SNPs within the *IL8* haplotypes. To do this, electrophoretic mobility shift assays (EMSA) were used to systematically screen for allele-specific differences in DNA-protein binding. When using A549 nuclear extracts, significant differences in protein-DNA binding were seen only at the +781 polymorphism. Inspection of the sequence surrounding the +781 SNP showed that a C/EBP β consensus binding site was present on the forward strand of the +781T allele but was disrupted on the +781C allele. A western blot showed that C/EBP β was present in A549 nuclear extract. In this same extract, supershift assays showed that C/EBP β binds to the +781T but not to the +781C polymorphism. The specific binding of C/EBP β to the +781T probe was confirmed using whole extracts from COS-7 cells previously transfected with a vector expressing C/EBP β. Intriguingly, therefore a transcription factor associated with the activation of genes central to the immune response appears to bind *in vitro* to the +781T allele found on the high producing *IL8* haplotype 2. These data suggest that the mechanism for *IL8* regulation may occur through differences in protein-DNA binding at functional polymorphisms within the *IL8* haplotype 2.

I investigated the nuclear extract of Human lymphocytic Jurkat cells in an attempt to define a mechanism for absence of allele-specific expression in primary lymphocytes. A western blot showed that C/EBP β was present in Jurkat cell nuclear extract. However, in supershift assays of Jurkat cell nuclear extract C/EBP β bound to neither the +781T allele nor the

+781C allele. It is possible, therefore that the cell specific nature of *IL8* regulation may occur through differences in C/EBP β binding.

The unusual mirror haplotype structure of *IL8* in European populations gave little indication of which SNP within the haplotype was likely to be functional *in vivo*. In order to demonstrate the principle that a non-coding polymorphism may alter transcriptional regulation *in vivo* I studied a second model system namely the tumor necrosis factor/lymphotoxin-α (*TNF/LTA*) locus with a novel technique. The haplotype-specific chromatin immunoprecipitation (haploChIP) technique developed in the Kwiatkowski laboratory to measure differences in the loading of RNA polymerase between alleles in EBV-immortalized B lymphocytes. The loading of RNA polymerase reflects the transcriptional activity of the gene and was used to define allele-specific gene expression in the absence of a coding polymorphism. It was this critical attribute that made it a suitable method with which to dissect the extent to which non-coding polymorphisms might modulate *IL8* regulation. Optimisation of the HaploChIP technique exploited the diverse haplotype structure of the *TNF/LTA* locus and the knowledge that allele-specific expression occurred on *LTA* induction to test for the presence of functional polymorphisms. A previous study has demonstrated that the *LTA*+80A allele is associated with low protein production through the recruitment of the transcriptional repressor protein Activated B cell Factor-1 (ABF-1). I extended this observation by demonstrating that allele-specific RNA polymerase loading was independently associated with the *LTA*+80, *LTA*+368 and *LTA*+723 polymorphisms. This association was refined to demonstrate that the polymerase loading at this locus appeared to be determined *in vivo* by the *LTA*+80A and *LTA*+368C alleles both of which occur on the same haplotype. These data may imply that

functional SNPs act in concert to influence transcription. Moreover, I infer that the functional basis for allele specific effects at critical candidate genes will have to be dissected *in vivo* at the level of the haplotype. HaploChIP has therefore shown itself to be a suitable method with which to understand the extent to which polymorphisms from within the *IL8* locus might alter transcription at that gene.

In summary therefore, this thesis provides evidence *in vivo* for haplotype-specific expression at the *IL8* locus that may explain the functional basis of the observed genetic association with RSV-bronchiolitis disease susceptibility. *In vitro* studies suggest that allele-specific binding of C/EBP β may be a factor in the *in vivo* haplotype-specific expression seen in primary respiratory cells. In a different *in vivo* model system I demonstrated that functional polymorphisms acting from within a haplotype can influence the degree to which RNA polymerase binds to an allele. Taken together these findings support the conclusion that genetic polymorphisms in non-coding DNA may influence the transcriptional regulation of genes important to the host's susceptibility to common infectious diseases.

APPENDIX 1

SYNERGISTIC INDUCTION OF NITRIC OXIDE BY CYTOKINES AND RSV

INTRODUCTION

A1.1.1

The role of nitric oxide in lung inflammation

Nitric Oxide (NO) is a highly reactive biological mediator with a broad range of roles which include smooth muscle relaxation, neurotransmission, inflammatory and immunosuppressive effects as well as the modulation of cytokine and chemokine production (Bogdan 2001). There is increasing evidence that nitric oxide (NO) plays a key role in both the physiological and pathological processes in the airways of the lung (Barnes 1995).

NO is produced from conversion of L-arginine and molecular oxygen to citrulline and NO by NO synthase (NOS). There are three types of NOS, neuronal NOS (NOS1), inducible NOS (NOS2) and endothelial NOS (NOS3). NOS1 and NOS3 are known collectively as constitutive NOS because unlike NOS2 they exist as constitutively expressed proteins which are regulated by calcium fluxes and calmodulin (Nathan et al. 1994) as well as at the level of transcription (Forstermann et al. 1998). Constitutively expressed NO is thought to act as a bronchodilator, a potent vasodilator of pulmonary circulation of the lung and may have some role in mucociliary clearance (Jain et al. 1993). Its expression is not restricted to neurons and endothelial cells. In contrast NOS2 expression is regulated by cytokines. For instance, human type II alveolar epithelial cell lines (A549) express NOS2 and produce NO when stimulated with a combination of IL-1β, TNFα and IFNγ (Asano et al. 1994). NOS2 expression is determined by *de novo* synthesis and stability of NOS2 mRNA and protein. In

addition, NOS2 produces NO at a high amplitude and is thought to function principally in response to infection. However, recent evidence suggests that all three isoforms of NOS have a role in the immune system (Bogden 2001).

A1.1.2

Role of inducible nitric oxide synthase in infectious disease

NO has a wide variety of actions in infection. NO has direct effects on microbes through the formation of free radicals, such as peroxynitrite ($ONOO^-$), which peroxidate membrane lipids, inhibit protein synthesis and mutate DNA. However, NO has a number of indirect functions such as cytokine regulation which may play an important role in the host response to infection. For instance, of relevance to the marked leukocyte response seen in RSV-induced bronchiolitis is the way in which NO can inhibit the activity of IL-8 (Sato et al. 2000). However, the expression of NOS2 in some infections is associated with a severe disease outcome through mechanisms such as NO-mediated cytotoxicity, induction of T cell apoptosis or from the generation of viral escape mutants (Akaike et al. 2000).

A1.1.3

Role of inducible nitric oxide from A549 cells in Respiratory syncytial virus infection

Recent reports have shown NO production from human A549 cells in response to stimulation with unpurified respiratory syncytial virus (RSV) both with (Tsutsumi et al. 1999) and without the addition of exogenous cytokines (Kao et al. 2001). However, RSV infected A549 cells produce IL-8, interleukin-1α (IL-1α), interleukin-1β (IL-1β) and tumor necrosis factor α (TNFα) (Patel et al. 1995, Patel et al. 1998). It is not clear from these experiments whether production of NO is due to the presence of contaminating cytokines in the cellular component of the viral preparation or whether the virus itself can directly induce NO.

RESULTS

A1.2.1

Synergistic induction of Nitric Oxide by Cytokines and RSV

RNI concentration was assayed in the supernatant of cells cultured in media alone (control), cells infected with purified RSV alone, cells treated with cytokines and from cells infected with purified RSV which were also treated with exogenous cytokines. As shown in Figure A.1, A549 cells infected with purified RSV produced little NO. When incubated with cytokines they produced significantly more RNI than uninfected cells incubated with cytokines alone (mean RNI 28.5 μM (SE 1.9) and 21.3 μM (SE 2.1) respectively, $p = 0.025$). Adjustment of the RNI for protein content, as shown in Figure A.2, demonstrates that A549 cells infected with purified virus alone do produce significantly more NO than controls (111.3 nM/μg (SE 25.2), 65.3 nM/μg (SE 4.9) respectively, $p = 0.01$). There was no

significant RNI production in cells exposed to inactivated RSV (data not shown). RNI production was inhibited with a 1 mM concentration of L-NMMA (data not shown).

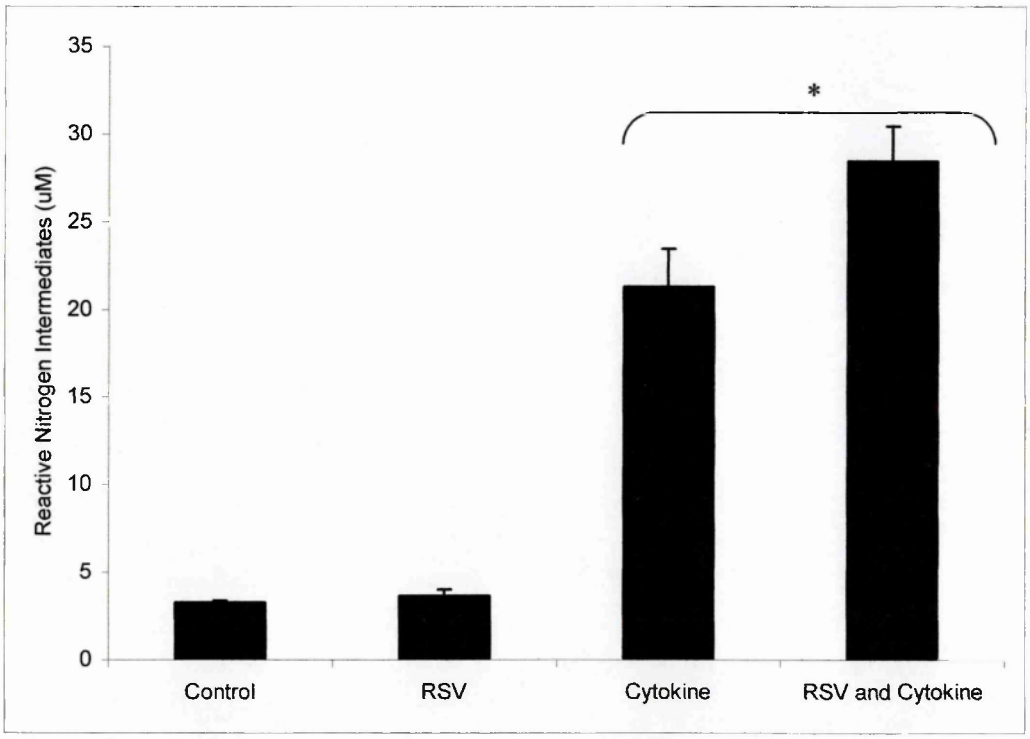


Figure A1.1. Mean Reactive Nitrogen Intermediates in A549 cell supernatants exposed to purified RSV, Cytokines and RSV with Cytokines for 72 hours. Vertical bars represent the standard errors of the mean for duplicate readings from five separate experiments.

* p = 0.025

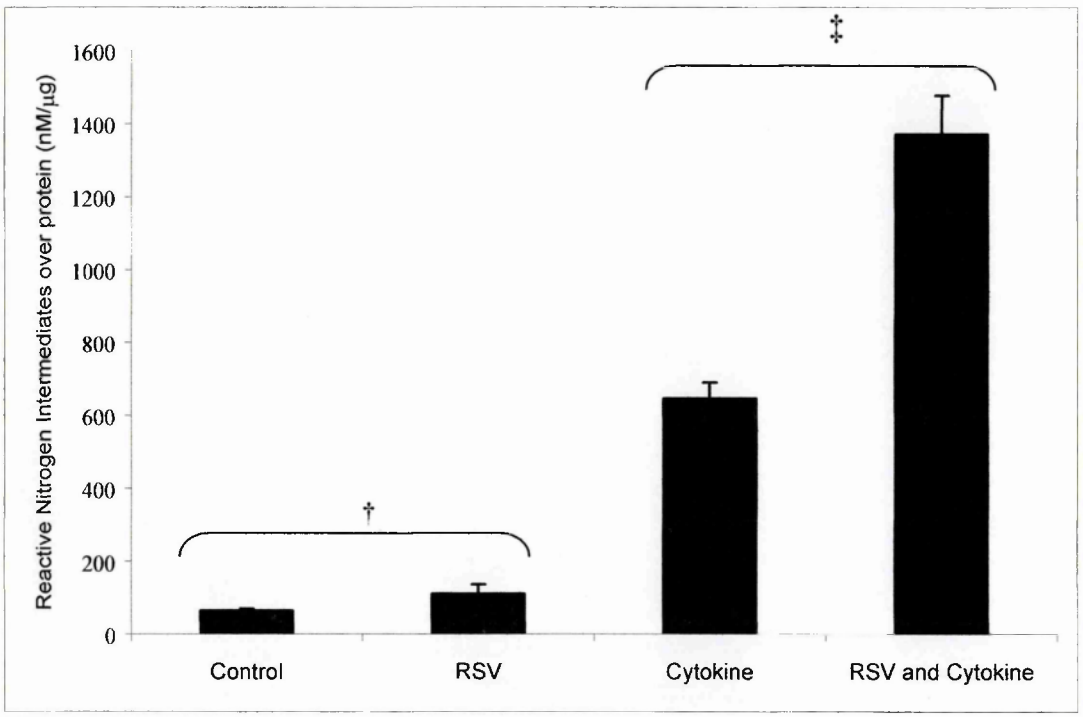


Figure A1.2. Mean Reactive Nitrogen Intermediates corrected for protein content in A549 cell supernatants exposed to purified RSV, Cytokines and RSV with Cytokines for 72 hours. Vertical bars represent the standard errors of the mean for duplicate readings from two separate experiments. †p = 0.01, ‡p = <0.001

DISCUSSION

A1.3.1

We have shown that infection with purified RSV produces relatively little NO from A549 cells at 72 hours in the absence of exogenous cytokines, despite evidence of viral protein expression at 24 hours and syncytia formation thereafter. This is in contrast to a previous report of relatively high levels of NO production when the same cell line is stimulated with unpurified viral preparations (Kao et al. 2001). This may be due to the removal during virus purification of known NO-activating cytokines such as TNFα and IL-1β both of which are present in unpurified RSV preparations grown on HEp2 cells within our laboratory□

When exogenous cytokines were added to A549 cells however, we observed that NO production was substantially increased by infection with purified virus. Synergistic induction of NO production has previously been demonstrated for combinations of cytokines (Asano et al. 1994), for hepatitis B virus with cytokines (Majono et al. 1998) and for unpurified RSV with cytokines (Tsutsumi et al 1999). These data suggest that the ability of RSV to induce NO production by respiratory epithelial cells may depend more on a cytokine cascade than on a direct stimulation effect of the virus.

APPENDIX 2

PUBLICATIONS

Hacking D, Rockett K., Hull J., Kwiatkowski D (2002). Synergistic action of cytokines and purified respiratory syncytial virus in Nitric Oxide induction. Journal Leukocyte Biology; 71: 729-730.

Hacking D, Hull J (2002). Respiratory syncytial virus – viral biology and the host response. Journal of Infection; 45: 1: 18-24.

Hacking D, Knight JC, Rockett K, Brown H, Frampton J, Kwiatkowski DP, Hull J and Udalova IA (2004). Increased in vivo transcription of an *IL8* haplotype associated with Respiratory Syncytial Virus disease-susceptibility. Genes and Immunity; 5: 274-282.

Acknowledgements

Acknowledgement of the assistance of colleagues who have contributed data from experiments they have conducted is given in Chapter 2 in the Methods section. In addition to the statements made in that section of the thesis I wish to thank, in alphabetical order, the following colleagues:

Helen Fox for the permission to include the Northern Blot of *IL8* after TNF stimulation (Chapter 3 , Figure 3.4) for preparing nuclear extract from A549 cells after TNF stimulation which was used in Chapter 6, Figure 6.1a.

Jeremy Hull for laying down the background work on which much of this thesis was based (as discussed in Chapter 1), his supervision of the thesis and his specific advice on the techniques of RSV purification, ELISAs, mRNA extraction, cDNA synthesis and MAPREC. I am also grateful to Jeremy for his help and advice in the drafting of Figures 1.1 to 1.4 in Chapter 1 for permission to reproduce Table 1.1 and Figure 1.5 (Chapter 1).

Brendan Keating for culturing and cross-linking the EBV immortalised CEPH B cell lines used in experiments in Chapter 7. I would also like to thank Brendan for his assistance with the DNA extraction phase in the haploChIP protocol.

Julian Knight for critical advice and more particularly on the planning and execution of allele-specific and haplotype-specific experiments. I am also grateful to Julian for inviting me to take part in the haploChIP project that he led and for tuition and supervision in

that investigation. I am grateful also for permission to reproduce Figures 7.1 and 7.2 in Chapter 7.

Dominic Kwiatkowski for inviting me to join the Childhood Infection Group which he leads, for his assistance in writing the grant which was funded by the MRC, the provision of further funds in addition to the MRC Fellowship grant and for his careful supervision of the thesis. I am grateful to Dominic his support and career advice over the last five years.

Medical Research Council of the United Kingdom for generous support in the form of an MRC Clinical Research Training Fellowship and the extension of that Fellowship over the months of September, October, November and December in 2003.

Kirk Rockett for his help in the drafting of the initial funding grants for the MRC, his critical advice throughout the thesis, his assistance with logistical support over the last three years and for specific technical tuition in the RNI assays seen in Appendix 1. I am grateful to Kirk for the writing of software used in the analysis of data from ELISAs, protein assays, RNI assays, and PE/MS quantitation experiments.

Irina Udalova for her critical advice throughout the thesis and specifically for tuition in the techniques of Northern Blotting, *in vitro* transfection experiments, UV-cross linking Western blots and EMSAs. I am grateful to Irina for permission to adapt one of her Figures shown in Chapter 5, Figure 5.3 and for the gift of the *IL8* gene expression construct shown in Chapter 5, Figure 5.1.

REFERENCES

Aherne W, Bird T, Court P, Gardner P, McQuillin J (1970). Pathological changes in virus infections of the lower respiratory tract in children. J Clin Pathol; 23: 7-18.

Akaike T, Fujii S, Kato A, Yoshitake J, Miyamoto Y, Sawa T, Okamoto S, Suga M, Asakawa M, Nagai Y, Maeda H. (2000). Viral mutation accelerated by nitric oxide production during infection *in vivo*. FASEB J; 10: 1447-1454.

Akira S, Kishimoto T (1997). NF-IL6 and NF-kappa B in cytokine gene regulation. Adv Immunol; 65: 1-46.

Albuquerque RV, Hayden CM, Palmer LJ, Laing IA, Rye PJ, Gibson NA, Burton PR, Goldblatt J, Lesouef PN (1998). Association of polymorphisms within the tumour necrosis factor (TNF) genes and childhood asthma. Clin. Exp. Allergy; 28: 578-584.

Alwan WH, Record FM, Openshaw PJ (1992). CD4+ T cells clear virus but augment disease in mice infected with respiratory syncytial virus. Comparison with the effects of CD8+ T cells. Clin Exp Immunol; 88: 527-536.

Amexis G, Oeth P, Abel K, Ivshina A, Pelloquin F, Cantor CR, Braun A, Chumakov K, Brau A. (2001) Quantitative mutant analysis of viral quasispecies by chip-based MALDI-TOF mass spectrometry. Proc Natl Acad Sci U S A.; 98: 12097-102. Epub 2001 Oct 02.

Arenberg DA, Kunkel SL, Polverini PJ, Glass M, Burdick MD, Strieter RM. (1996) Inhibition of interleukin-8 reduces tumorigenesis of human non-small cell lung cancer in SCID mice. J Clin Invest; 97:2792-802.

Asano K, Chee CB, Gaston B, Lilly CM, Gerard C, Drazen JM, Stamler JS (1994). Constitutive and inducible nitric oxide synthase gene expression, regulation, and activity in human lung epithelial cells. Proc Natl Acad Sci USA; 91: 10089-10093.

Badenhoop K, Schwarz G, Schleusener H, Weetman AP, Recks S, Peters H, Bottazzo GF, Usadel KH (1992). Tumor necrosis factor beta polymorphisms in Graves' disease. J. Clin. Endocrinol. Metab. 74, 287-291.

Barnes PJ (1995). Nitric Oxide and airway disease. Ann Med; 27: 389-393.

Barr FE, Pedigo H, Johnson TR, Shepherd VL (2000). Surfactant protein-A enhances uptake of respiratory syncytial virus by monocytes and U937 macrophages. Am J Respir Cell Mol Biol; 23: 586-592.

Becker S, Quay J, Soukup J (1991). Cytokine (tumor necrosis factor, IL-6, and IL-8) production by respiratory syncytial virus-infected human alveolar macrophages. J Immunol; 147: 4307-4312.

Becker S, Reed W, Henderson FW, Noah TL (1997). RSV infection of human airway epithelial cells causes production of the B-chemokine RANTES. Am J Physiol; 272: L512-L520.

Bioque G, Crusius JB, Koutroubakis I, Bouma G, Kostense PJ, Meuwissen SG, Pena AS (1995). Allelic polymorphism in IL-1 beta and IL-1 receptor antagonist (IL-1Ra) genes in inflammatory bowel disease. Clin Exp Immunol; 102: 379-383.

Bogdan C (2001). Nitric Oxide and the immune response. Nature Immunology; 2: 907-916.

Bray MS, Boerwinkle E, Doris PA (2001). High-throughput multiplex SNP genotyping with MALDI-TOF mass spectrometry: practice, problems and promise. Hum Mutat 2001; 17, 296-304.

Bray NJ, Buckland PR, Owen MJ, O'Donovan MC (2003a). Cis-acting variation in the expression of a high proportion of genes in human brain. Hum Genet; 113: 149-153.

Bray NJ, Buckland PR, Williams NM, Norten N, Owen MJ, O'Donovan MC (2003b) A haplotype implicated in schizophrenia susceptibility is associated with reduced COMT expression in human brain. Am J Hum Genet; 73: 152-161.

Bristos-Bray M, and Friedman AD (1997). Core binding factor cannot synergistically activate the myeloperoxidase proximal enhancer in immature myeloid cells without c-Myb. Mol. Cell. Biol; 17: 5127-5135.

Cane PA (1997). Analysis of linear epitopes recognized by the primary antibody response to variable region of the attachment (G) protein of respiratory syncytial virus. J Med Virol; 51: 297-304.

Cane PA (2001). Molecular epidemiology of respiratory syncytial virus. Rev Med Virol; 11: 103-116.

Casola A, Garofalo RP, Jamaluddin M, Vlahopolos S, Brasier AR (2000). Requirement of a novel upstream response element in respiratory syncytial virus-induced *IL8* gene expression. J Immunol; 164: 5944-5951.

Chandwani S, Borkowsky W, Krasinski K, Lawrence R, Welliver R (1990). Respiratory syncytial virus infection in human immunodeficiency virus-infected children. J Pediatr; 117: 251-254.

Chanock RM, Roizman B, Myers R (1957). Recovery from infants with respiratory illness of a virus related to chimpanzee coryzal agent (CCA). I. Isolation, properties and characterization. Am J Hyg; 66: 281-290.

Cheung VG, Conlin LK, Weber TM, Arcaro M, Jen K, Morley M, Spielman RS (2003) Natural Variation in human gene expression assessed in lymphoblastoid cells. Nat Genet; 33: 422-425.

Cherrie AH, Anderson K, Wertz GW, Openshaw PJ (1992). Human cytotoxic T cells stimulated by antigen on dendritic cells recognize the N, SH, F, M, 22K, and 1b proteins of respiratory syncytial virus. J Virol; 66: 2102-2110.

Chiba Y, Higashidate Y, Suga K, Honjo K, Tsutsumi H, Ogra PL (1989). Development of cell-mediated cytotoxic immunity to respiratory syncytial virus in human infants following naturally acquired infection. J Med Virol; 28: 133-139.

Cho EJ, Kobor MS, Kim M, Greenblatt J, Buratowski S (2001). Opposing effects of Ctk1 kinase and Fcp1 phosphatase at Ser 2 of the RNA polymerase II C-terminal domain. Genes Dev; 15: 3319-3329.

Collins PL, Mottet G (1991). Post-translational processing and oligomerization of the fusion glycoprotein of human respiratory syncytial virus. J Gen Virol; 72: 3095-3101.

Collins PL, McIntosh K, Chanock RM (1996). Respiratory syncytial virus. In: Fields BN, Knipe DM, Howley PM *et al,* eds. Fields Virology. Lippincott-Raven: 1313-1343.

Cowles CR, Hirschhorn JN, Altshuler D, Lander ES (2002). Detection of regulatory variation in mouse genes. Nat Genet; 32: 432-437.

Enard W, Khaitovich P, Klose J, Zollner S, Heissig F, Giavalisco P, Nieselt-Struwe K, Muchmore E, Varki A, Ravid R, Doxiadis GM, Bontrop RE, Paabo S. (2002) Intra- and interspecific variation in primate gene expression patterns. Science; 296: 340-343.

Ess K, Witte D, Bascomb C, Aronow B (1999). Diverse developing mouse lineages exhibit high-level c-Myb expression in immature cells and loss of expression upon differentiation. Oncogene; 18: 1103-1111.

Everard ML, Swarbuck A, Wrightham M, McIntyre J, Dunkley C, James PD, Sewell HF, Miller AD (1994). Analysis of cells obtained by bronchial lavage of infants with respiratory syncytial virus infection. Arch Dis Child; 71: 428-432.

Feldman SA, Hendry RM, Beeler JA (1999). Identification of a linear heparin binding domain for human respiratory syncytial virus attachment glycoprotein G. J Virol; 73: 6610-6617.

Fiedler MA, Wernke-Dollries K, Stark JM. (1995). Respiratory syncytial virus increases *IL8* gene expression and protein release in A549 cells. Am J Physiol; 269: L865-L872.

Fiedler MA, Wernke-Dollries K, Stark JM. (1996). Mechanism of RSV-induced *IL8* gene expression in A549 cells before viral replication. Am J Physiol; 271: L963-L971.

Fiedler MA, Wernke-Dollries K, Stark JM (1998). Inhibition of TNF-alpha-induced NF-kappaB activation and IL-8 release in A549 cells with the proteasome inhibitor MG-132. Am J Respir Cell Mol Biol; 19:259-68.

Fletcher JN, Smyth RL, Thomas HW, Ashby D, Hart A (1997). Respiratory syncytial virus genotypes and disease among children in hospital. Arch Dis Child; **77**: 508-511.

Fletcher, C, Heintz, N, Roeder, RG (1987). Purification and characterization of OTF-1, a transcription factor regulating cell cycle expression of a human histone H2b gene. Cell; 51: 773-81.

Forstermann U, Boissel J, Kleinert H (1998). Expressional control of the 'constitutive' isoforms of nitric oxide synthase (NOS I and NOS III). FASEB J; 12: 773-790.

Gambaro G, Anglani F, D'Angelo A (2000) Association studies of genetic polymorphisms and complex disease. Lancet; 355: 308-11.

Garofalo R, Kimpen JL, Welliver RC, Ogra PL (1992). Eosinophil degranulation in the respiratory tract during naturally acquired respiratory syncytial virus infection. J Pediatr; 120: 28-32.

Garofalo R, Sabry M, Jamaluddin M, Yu RK, Casola A, Ogra PL, Brasier AR (1996). Transcriptional activation of the interleukin-8 gene by respiratory syncytial virus infectionin alveolar epithelial cells: Nuclear translocation of the RelA transcription factor as a mechanism producing airway mucosal inflammation. J Virol; 70: 8773-8781.

Garofalo RP, Haeberle H (2000). Epithelial regulation of innate immunity to respiratory syncytial virus. Am J Respir Cell Mol Biol; 23: 581-5.

Garofalo R, Patti J, Hintz KA, Hill V, Ogra P, Welliver R (2001). Macrophage inflammatory protein-1a (not T helper type 2 cytokines) is associated with severe forms of respiratory syncytial virus bronchiolitis. J Infect Dis; 184: 393-9.

Gerard C and Rollins BJ (2001). Chemokines and disease. Nature Immunol; 2: 108-115.

Gerszten RE, Garcia-Zepeda EA, Lim YC, Yoshida M, Ding HA, Gimbrone MA Jr, Luster AD, Luscinskas FW, Rosenzweig A (1999). MCP-1 and IL-8 trigger firm adhesion of monocytes to vascular endothelium under flow conditions. Nature; 398: 718-23.

Ghildyal R, Hartley C, Varrasso A, Meanger J, Voelker DR, Anders EM, Mills J (1999). Surfactant protein A binds to the fusion glycoprotein of respiratory syncytial virus and neutralizes virion infectivity. J Infect Dis; 180: 2009-2013.

Giard DJ, Aaronson SA, Todaro GJ, Arnstein P, Kersey JH, Dosik H, Parks WP (1973). *In vitro* cultivation of human tumors: establishment of cell lines derived from a series of solid tumors. J Natl Cancer Inst; 51: 1417-1423.

Gower TL, Graham BS (2001). Antiviral activity of lovastatin against respiratory syncytial virus in vivo and in vitro. Antimicrob Agents Chemother; 45: 1231-1237.

Gower TL, Peeples ME, Collins PL, Graham BS (2001). RhoA is activated during respiratory syncytial virus infection. Virology; 283: 188-96.

Graf T (1992). Myb: a transcriptional activator linking proliferation and differentiation in hematopoietic cells. Curr Opin Genet Dev; 2: 249-55.

Graham BS, Henderson GS, Tang YW, Lu X, Neuzil KM, Colley DG (1993). Priming immunization determines T helper cytokine mRNA expression patterns in lungs of mice challenged with respiratory syncytial virus. J Immunol; 151: 2032-2040.

Graham BS (1995). Pathogenesis of respiratory syncytial virus vaccine-augmented pathology. Am J Respir Crit Care Med; 152: S63-66.

Haynes LM, Moore DD, Kurt-Jones EA, Finberg RW, Anderson LJ, Tripp RA (2001). Involvement of toll-like receptor 4 in innate immunity to respiratory syncytial virus. J Virol; 75:10730-10707.

Holberg CJ, Wright AL, Martinez FD, Ray CG, Taussig LM, Lebowitz MD (1991). Risk factors for respiratory syncytial virus-associated lower respiratory illnesses in the first year of life. Am J Epidemiol; 133: 1135-1151.

Holmes WE, Lee J, Kuang WJ, Rice GC, Wood WI (1991). Structure and functional expression of a human interleukin-8 receptor. Science; 253: 1278-1280.

Hull J, Thomson A, Kwiatkowski D (2000). Association of respiratory syncytial virus bronchiolitis with the interleukin 8 gene region in UK families. Thorax; 55: 1023-1027.

Hull J, Ackerman H, Isles K, Usen S, Pinder M, Thomson A, Kwiatkowski D (2001). Unusual haplotypic structure of IL8, a susceptibility locus for a common respiratory virus. Am J Hum Genet; 69: 413-9.

Hull J, Rowlands K, Lockhart E, Sharland M, Moore C, Hanchard N, Kwiatkowski D (2004). Haplotype mapping of the bronchiolitis susceptibility locus near *IL8*. Hum Genet; 114: 272-279.

IMpact-RSV Study Group (1998). Palivizumab, a humanized respiratory syncytial virus monoclonal antibody, reduces hospitalization from respiratory syncytial virus infection in high-risk infants. Pediatrics; 102: 531-537.

Ioannidis JP, Ntzani EE, Trikalinos TA, Contopoulos-Ioannidis DG (2001). Replication validity of genetic association studies. Nat Genet; 29: 306-9.

Jain B, Lubinstein I, Robbins RA, Leise KL, Sisson JH (1993). Modulation of airway epithelial cell ciliary beat frequency by nitric oxide. Biochem Biophys Res Commun; 191: 83-88.

Jamaluddin M, Garofalo R, Ogra PL, Brasier AR (1996). Inducible translational regulation of the NF-IL6 transcription factor by respiratory syncytial virus infection in pulmonary epithelial cells. J Virol; 70: 1554-1563.

Jamaluddin MA, Casola A, Garofalo RP, Han Y, Elliot T, Orga PL, Brasier AR. (1998). A major component of IκBα proteolysis occurs independently of the proteosme pathway in respiratory syncytial virus-infected pulmonary epithelial cells. J Virol; 72; 6: 4849-4857.

Jiang ZD, Okhuysen PC, Guo DC, He R, King TM, Dupont HL, Milewicz DM (2003). Genetic susceptibility to enteroaggregative Escherichia coli diarrhea: polymorphism in the interleukin-8 promotor region J Infect Dis; 188: 506-11.

Johnson PR, Spriggs MK, Olmsted RA, Collins PL (1987). The G glycoprotein of human respiratory syncytial virus of subgroups A and B: extensive sequence divergence between antigenically related proteins. Proc Natl Acad Sci USA; 84: 5625-5629.

Johnson NA and Porter AH (2000). Rapid speciation via parallel, directional selection on regulatory genetic pathways. J Theor Biol; 205: 527-542.

Kao YJ, Piedra PA, Larsen GL, Colasurdo GN (2001). Induction and regulation of Nitric oxide synthase in airway epithelial cells by Respiratory syncitial virus Am J Respir Crit Care Med; 163: 532-539.

Karron RA, Buonagurio DA, Georgiu AF et al (1997). Respiratory syncytial virus (RSV) SH and G proteins are not essential for viral replication in vitro: Clinical evaluation and molecular characterization of a cold passaged, attenuated RSV subgroup B mutant. Proc Natl Acad Sci USA; 94: 13961-13966.

Kim HW, Canchola JG, Brandt CD (1969). Respiratory syncytial virus disease in infants despite prior administration of antigenic inactivated vaccine. Am J Epidemiol; 89: 422-434.

Knight JC, Udalova I, Hill AV, Greenwood BM, Peshu N, Marsh K, Kwiatkowski D (1999). A polymorphism that affects OCT-1 binding to the TNF promoter region is associated with severe malaria. Nat Genet; 22: 145-50.

Knight JC, Keating BJ, Rockett KA, Kwiatkowski DP (2003). *In vivo* characterization of regulatory polymorphisms by allele-specific quantification of RNA polymerase loading. Nat Genet; 33: 469-475.

Knight JC (1994). Allele-specific gene expression uncovered. Trends in Genetics; *In press*

Knight JC, Keating BJ, Kwiatkowski DP (2004). Allele-specific repression of lymphotoxin-alpha by Activated B cell Factor-1. Nat Genet; *In press.*

Koch AE, Polverini PJ, Kunkel SL, Harlow LA, DiPietro LA, Elner VM, Elner SG, Strieter RM (1992). Interleukin-8 as a macrophage-derived mediator of angiogenesis. Science; 258: 1798-1801.

Komarnitsky P, Cho EJ, Buratowski S (2000). Different phosphorylated forms of RNA polymerase II and associated mRNA processing factors during transcription. Genes Dev; 14: 2452-2460.

Kowenz-Leutz E, Herr P, Niss K, Leutz A (1997). The homeobox gene *GBX2*, a target of the *myb* oncogene, mediates autocrine growth and monocyte differentiation. Cell; 91: 185-195.

Krug N, Madden J, Redington AE, Lackie P, Djukanovic R, Schauer U, Holgate ST, Frew AJ, Howarth PH (1996). T-cell cytokine profile evaluated at the single cell level in BAL and blood in allergic asthma. Am J Respir Cell Mol Biol.; 14: 319-326.

Kurt-Jones EA, Popova L, Kwinn L, Haynes LM, Jones LP, Tripp RA, Walsh EE, Freeman MW, Golenbock DT, Anderson LJ, Finberg RW (2000). Pattern recognition receptors TLR4 and CD14 mediate response to respiratory syncytial virus. Nat Immunol; 1: 398-401.

Lambert DM, Barney S, Lambert AL et al (1996). Peptides from conserved regions of paramyxovirus fusion (F) proteins are potent inhibitors of viral fusion. Proc Natl Acad Sci U S A; 93: 2186-2191.

LeVine AM, Gwozdz J, Stark J, Bruno M, Whitsett J, Korfhagen T (1999). Surfactant protein-A enhances respiratory syncytial virus clearance in vivo. J Clin Invest; 103: 1015-1021.

Levine S, Klaiber-Franco R, Pradiso PR (1987). Demonstration that glycoprotein G is the attachment protein of respiratory syncitial virus. J Gen Virol; 68: 2521-2524.

Levine M (2002) How insects lose their limbs Nature; 415: 848- 849.

Lo HS, Wang Z, Hu Y, Yang H, Gere S, Buetow K, Lee M (2003). Allelic variation in gene expression is common in the human genome. Genome Res; 13:1855-62.

Ma X, Reich RA, Wright JA, Tooker HR, Teeter LD, Musser JM, Graviss ED (2003) Association between interleukin-8 gene alleles and human susceptibility to tuberculosis disease. J Infect Dis; 188: 349-55.

MacKay CR (2001). Chemokines: immunology's high impact factors. Nat Immunol; 2: 95-101.

Majetschak M, Flohe S, Obertacke U, Schroder J, Staubach K, Nast-Kolb D, Schade FU, Stuber F (1999). Relation of a TNF gene polymorphism to severe sepsis in trauma patients. Ann Surg; 230: 207-14.

Majono PL, Garcia-Monzon C, Lopez-Cabrera M, Lara-Pezzi E, Fernandez-Ruiz E, Garcia-Iglesias C, Borque MJ, Moreno-Otero R (1998). Inducible Nitric Oxide Synthase expression in chronic viral hepatitis. J Clin Invest; 101: 1343-1352.

Manak JR, Mitiku N, Lipsick JS (2002). Mutation of the Drosophila homologue of the Myb protooncogene causes genomic instability. Proc Natl Acad Sci U S A.; 99: 7438-43.

Matsusaka T, Fujikawa K, Nishio Y et al. (1993). Transcription factors NF-IL6 and NF-kappa B synergistically activate transcription of the inflammatory cytokines, interleukin 6 and interleukin 8. Proc Natl Acad Sci USA; 90: 10193-10197.

Mastronarde JG, He B, Monick MM, Mukaida N, Matsushima K, Hunninghake GW (1996). Induction of interleukin (IL)-8 gene expression by respiratory syncytial virus involves activation of nuclear factor (NF)-kappa B and NF-IL-6. J Infect Dis; 174: 262-7.

Mastronade JG, Monick MM, Mukaida N, Matsushima K, Hunninghake GW (1998). Activator protein-1 is the preferred transcription factor for cooperative interaction with Nuclear Factor-κB in respiratory syncytial virus-induced interleukin-8 gene expression in air way epithelium. J Infect Dis; 177: 1275-1281.

Miau LH, Chang CJ, Shen BJ, Tsai WH, Lee SC (1998). Identification of heterogeneous nuclear ribonucleoprotein K (hnRNP K) as a repressor of C/EBPbeta-mediated gene activation. J Biol Chem; 273: 10784-10791.

McCarron SL, Edwards S, Evans PR, Gibbs R, Dearnaley DP, Dowe A, Southgate C, Easton DF, Eeles RA, Howell WM (2002). Influence of cytokine gene polymorphisms on the development of prostate cancer. Cancer Res; 62: 3369-72.

Medical Research Council Subcommittee report on Respiratory Syncytial Virus Vaccines (1978). Respiratory syncytial virus infection: admissions to hospital in industrial, urban, and rural areas. Br Med J; 2: 796-798.

Midulla F, Villani A, Panuska JR, Dab I, Kolls JK, Merolla R, Ronchetti R (1993). Respiratory syncytial virus lung infection in infants: immunoregulatory role of infected alveolar macrophages. J Infect Dis; 168: 1515-1519.

Modi WS, Chen ZQ (1998). Localization of the human CXC chemokine subfamily on the long arm of chromosome 4 using radiation hybrids. Genomics; 47: 136-9.

Modi WS and Yoshimura T (1999). Isolation of novel GRO genes and a phylogenetic analysis of the CXC chemokine subfamily in mammals. Mol. Biol. Evol; 16: 180-193.

Moffatt, M.F. & Cookson, W.O (1997). Tumour necrosis factor haplotypes and asthma. Hum. Mol. Genet. 6, 551-554.

Mosmann TR, Coffman RL (1989). TH1 and TH2 cells: different patterns of lymphokine secretion lead to different functional properties. Annu Rev Immunol; 7: 145-173.

Murphy PM. (2001). Viral expoitation and subversion of the immune system through chemokine mimicry. Nature Immunology; 2: 116-122.

Nathan C and Xie Q (1994). Nitric Oxide synthase: roles, tolls and controls. Cell; 78; 915-918.

Newton R, Adcock IM and Barnes, PJ (1996) Superinduction of NF-kappa B by actinomycin D and cycloheximide in epithelial cells. Biochem Biophys Res Commun; 218: 518-23.

Ness SA, Kowenz-Leutz E, Casini T, Graf T and Leutz A (1993). Myb and NF-M: combinatorial activators of myeloid genes in heterologous cell types. Genes Dev; 7: 749-759.

Noah TL, Becker S (1993). Respiratory syncytial virus-induced cytokine production by a human bronchial epithelial cell line. Am J Physiol; 265: L472-L478.

Noah TL, Henderson FW, Wortman IA, Devlin RB, Handy J, Koren HS, Becker S (1995). Nasal Cytokine production in viral acute upper respiratory infection of childhood. J Infect Dis; 171: 584-592.

Nourbakhsh M, Kailbe S, Dorrie A, Hauser H, Resch K, Kracht M. (2001). The NF-κB repressing factor is involved in basal repression and interleukin (IL)-1 induced activation of *IL8* transcription by binding to a conserved NF-κB-flanking sequence element. J Biol Chem; 276: 6: 4501-4508.

Oelgeschlager M, Janknecht R, Krieg J, Schreek S, and Luscher B (1996). Interaction of the coactivator CBP with Myb proteins: effects on Myb-specific transactivation and on the cooperativity with NF-M. EMBO J; 15: 2771-2780.

Oleksiak MF, Churchill GA, Crawford DL. (2002). Variation in gene expression within and among natural populations. Nat Genet; 32: 261-266.

Osada S, Yamamoto H, Nishihara T, Imagawa M (1996). DNA binding specificity of the CCAAT/enhancer-binding protein transcription factor family. J Biol Chem; 271: 3891-3896.

Ozaki K, Ohnishi Y, Iida A, Sekine A, Yamada R, Tsunoda T, Sato H, Sato H, Hori M, Nakamura Y, Tanaka T (2002). Functional SNPs in the lymphotoxin-alpha gene that are associated with susceptibility to myocardial infarction. Nat Genet; 32: 650-4. Epub 2002 Nov 11. Erratum in: Nat Genet. 2003; 33:107.

Palivizumab, a humanized respiratory syncytial virus monoclonal antibody, reduces hospitalization from respiratory syncytial virus infection in high-risk infants (1998) The IMpact-RSV Study Group. Pediatrics; 102: 531-537.

Patel JA, Kunimoto M, Sim TC, Garofalo R, Eliott T, Baron S, Ruuskanen O, Chonmaitree T, Ogra PL, Schmalstieg F (1995). Interleukin-1α mediates the enhanced expression of intracellular adhesion molecule-1 in pulmonary epithelial cells infected with respiratory syncytial virus. Am J Respir Cell Mol Biol; 13: 602-609.

Patel JA, Jiang Z, Nakajima N, Kunimoto M (1998). Autocrine regulation of interleukin-8 by interleukin-1α in respiratory syncytial virus-infected pulmonary epithelial cells in vitro. Immunology; 95: 501-506.

Pastey MK, Crowe JE, Graham BS (1999). RhoA interacts with the fusion glycoprotein of respiratory syncytial virus and facilitates virus-induced syncytium formation. J Virol; 73: 7262-7270.

Pastey MK, Gower TL, Spearman PW, Crowe JE Jr, Graham BS (2000). A RhoA-derived peptide inhibits syncytium formation induced by respiratory syncytial virus and parainfluenza virus type 3. Nat Med.; 6: 35-40.

Patthy A, Bajusz S, Patthy L (1977). Preparation and characterization of N^G-mono, di-, and trimethylated arginines. Acta Biochem. Biophys. Acad. Sci. Hung; 12: 191-195.

Pennacchio LA and Rubin EM (2001). Genomic strategies to identify mammalian regulatory sequences. Nat Rev Genet; 2: 100-9.

Petersen F, Ludwig A, Flad HD, Brandt E (1996). TNF-alpha renders human neutrophils responsive to platelet factor 4. Comparison of PF-4 and IL-8 reveals different activity profiles of the two chemokines. J Immunol; 156: 1954-62.

Poli V (1998). The role of C/EBP isoforms in the control of inflammatory and native immunity functions. J Biol Chem; 273: 29279-29282.

Quandt K, Frech K, Karas H, Wingender E, Werner T. MatInd and MatInspector: new fast and versatile tools for detection of consensus matches in nucleotide sequence data. Nucleic Acids Res; 23: 4878-84.

Rieder MJ, Taylor SL, Clark AG, Nickerson DA (1999). Sequence variation in the human angiotensin converting enzyme. Nat Genet; 22: 1: 59-62.

Risch N, Merikangas K (1996). The future of genetic studies of complex human diseases. Science; 273: 1516-1517.

Rockett KA, Brookes R, Udalova I, Vidal V, Hill AVS, Kwiatkowski D (1998). 1,25-Dihydroxyvitamin D_3 induces nitric oxide synthase and suppresses growth of Mycobacterium tuberculosis in a Human macrophage-like cell line. Infec Immun; 66: 11: 5314-5321.

Roger T, Out T, Mukaida N, Matsushima K, Jansen H, Lutter R. (1998) Enhanced AP-1 and NF-kappaB activities and stability of interleukin 8 (IL-8) transcripts are implicated in *IL8* mRNA superinduction in lung epithelial H292 cells. Biochem J; 330:429-35.

Roman M, Calhoun WJ, Hinton KL, Avendano LF, Simon V, Escobar AM, Gaggero A, Diaz PV (1997). Respiratory syncytial virus infection in infants is associated with predominant Th-2-like response. Am J Respir Crit Care Med.; 156: 190-195.

Rossi D, Zlotnik A (2001). The biology of chemokines and their receptors. Annu Rev Immunol; 18: 217-242.

Sala A and Watson R (1999). B-Myb protein in cellular proliferation, transcription control, and cancer: latest developments. J Cell Physiol; 179: 245-50.

Sato E, Simpson KL, Grisham MB, Koyama S, Robbins RA (2000). Reactive nitrogen and oxygen species attenuate interleukin-8-induced neutrophil chemotactic activity *in vitro*. J Biol Chem; 275: 10826-10830.

Schlender J, Bossert B, Buchholtz U, Conzelman KK (2000). Bovine respiratory syncytial virus nonstructural proteins NS1 and NS2 cooperatively antagonize α/β interferon-induced antiviral response. J Virol; 74: 8234-8242.

Schmidt AC, Couch RB, Galasso GJ, Hayden FG, Mills J, Murphy BR, Chanock RM. (2001). Current research on respiratory viral infections: Third International Symposium. Antiviral Research; 50: 157-196.

Schroeder SC, Schwer B, Shuman S, Bentley D (2000). Dynamic association of capping enzymes with transcribing RNA polymerase II. Genes Dev; 14: 2435-2440.

Sigma-Aldrich Technical Bulletin MB-205 (1999). TRI reagent Product No. T9424. In: Sigma Technical Bulletin MB-205. Sigma, Saint Louis, MS 63103; email: sigma-techserv@sial.com.

Schreiber E, Matthias P, Muller MM, Schaffner W (1989). Rapid detection of octamer binding proteins with 'mini-extracts', prepared from a small number of cells. Nucleic Acids Res; 17: 15: 6419.

Selwyn BJ (1990). The epidemiology of acute respiratory tract infection in young children: comparison of findings from several developing countries. Coordinated Data Group of BOSTID Researchers. Rev Infect Dis; 12 Suppl 8: S870-88.

Sitzmann J, Noben-Trauth K, Klempnauer K (1995). Expression of mouse *c-myb* during embryonic development. Oncogene; 11: 2273-2279.

Shay DK, Holman RC, Roosevelt GE, Clarke MJ and Anderson LJ (2001). Bronchiolitis – associated mortality and estimates of Respiratory syncytial virus-associated deaths among US children, 1979-1997. J Infect Dis; 183: 16-22.

Smith JA (1993). Antibdy-sandwich ELISA to detect soluable antigens. In: Current Protocols in Molecular Biology, edited by FM Ausubel, R Brent, RE Kingston, DD Moore, JG Seidman, JA Smith, and K Struhl. New York: Wiley, p. 11.2.8-11.2.9.

Smith DR, Polverini PJ, Kunkel SL, Orringer MB, Whyte RI, Burdick MD, Wilke CA, Strieter RM (1994). Inhibition of interleukin 8 attenuates angiogenesis in bronchogenic carcinoma. J Exp Med; 179: 1409-1415.

Smyth RL, Mobbs KJ, O'Hea U, Ashby D, Hart CA (2002). Respiratory syncytial virus bronchiolitis: disease severity, interleukin-8, and virus genotype. Pediatr Pulmonol; 33: 339-346.

Srikiatkhachorn A, Braciale TJ (1997). Virus-specific CD8+ T lymphocytes downregulate T helper cell type 2 cytokine secretion and pulmonaryeosinophilia during experimental murine respiratory syncytial virus infection. J Exp Med; 186: 421-432.

Stata Corporation (2001). Statistical functions. In: STATA statistical software user's guide. College Station, TX: Stata Press; Release 7.0: 113-119.

Stuber F, Petersen M, Bokelmann F and Schade, U (1996). A genomic polymorphism within the tumor necrosis factor locus influences plasma tumor necrosis factor-alpha concentrations and outcome of patients with severe sepsis. Crit Care Med; 24: 381-384.

Tabor HK, Risch NJ, Myers RM (2002). Opinion: Candidate-gene approaches for studying complex genetic traits: practical considerations. Nat Rev Genet; 3: 5: 391-7.

Tahirov TH, Sato K, Ichikawa-Iwata E, Sasaki M, Inoue-Bungo T, Shiina M, Kimura K, Takata S, Fujikawa A, Morii H, Kumasaka T, Yamamoto M, Ishii S, Ogata K (2002).

Mechanism of c-Myb-C/EBP beta cooperation from separated sites on a promoter. Cell; 108: 57-70.

Takai Y (1995). Rho as a regulator of the cytoskeleton. Trends Biochem Sci; 20: 227-231.

Takahashi Y, Rayman JB, Dynlacht BD (2000). Analysis of promoter binding by the E2F and pRB families in vivo: distinct E2F proteins mediate activation and repression. Genes Dev; 14: 804-16.

Taylor CE, Morrow S, Scott M, Young B, Toms GL (1989). Comparative virulence of respiratory syncitial virus subgroups A and B. Lancet; I: 777-778.

Teng MN, Collins PL (1998). Identification of the respiratory syncytial virus proteins required for fomation and passage of helper-dependent infectious particles. J Virol; 72: 5707-5716.

Teng MN, Whitehead SS, Collins PL (2001). Contribution of the respiratory syncytial virus G glycoprotein and its secreted and membrane-bound forms to virus replication in vitro and in vivo. Virology; 289: 283-296.

Terry CF, Loukaci V, and Green FR (2000). Cooperative influence of genetic polymorphisms on interleukin 6 transcriptional regulation. J. Biol. Chem; 275: 18138-44.

Tripp RA, Jones LP, Haynes LM, Zheng H, Murphy PM, Anderson LJ. (2001). CX3C chemokine mimicry by respiratory syncytial virus G glycoprotein. Nature Immunol; 2: 732-738.

Tsutsumi H, Takeuchi R, Ohsaki M, Seki K, Chiba S (1999). Respiratory syncytial virus infection of human respiratory epithelial cells enhances inducible nitric oxide synthase expression. J Leukoc Biol; 66: 99-104.

Tsuchihashi Z, Dracopoli NC (2002). Progress in high throughput SNP genotyping methods. Pharmacogenomics J; 2: 103-110.

Turner DM, Williams DM, Sankaran D, Lazarus M, Sinnott PJ, Hutchinson IV (1997). An investigation of polymorphism in the interleukin-10 gene promoter. Eur J Immunogenet; 1: 1-8.

Ueba O (1978). Respiratory syncitial virus. I. Concentration and purification of the infectious virus. Acta Med Okayama: 32: 265-272.

Udalova I, Knight J, Vidal V, Nedospasov S, Kwiatkowski D (1998). Complex NF-kappaB interactions at the distal tumor necrosis factor promoter region in human monocytes. J Biol Chem; 273: 21178-21186.

Udalova IA, Richardson A, Denys A, Smith C, Ackerman H, Foxwell B, Kwiatkowski D. (2000) Functional consequences of a polymorphism affecting NF-kappaB p50-p50 binding to the TNF promoter region. Mol Cell Biol; 20: 9113-9119.

Udalova IA, Mott R, Field D, Kwiatkowski D (2002). Quantitative prediction of NF-kappa B DNA-protein interactions. Proc Natl Acad Sci U S A.; 99: 8167-72.

van Schaik SM, Tristram DA, Nagpal IS, Hintz KM, Welliver RC 2nd, Welliver RC (1999). Increased production of IFN-gamma and cysteinyl leukotrienes in virus-induced wheezing. J Allergy Clin Immunol; 103: 630-636.

Wang D, Baldwin AS (1998). Activation of Nuclear Factor-kB-dependent transcription by Tumour necrosis factor-α is mediated through phosphylation of Rel/p65 on serine 529. J Biol Chem; 273: 45: 29411-29416.

Wang SZ, Xu H, Wraith A, Bowden JJ, Alpers JH, Forsyth KD (1998). Neutrophils induce damage to respiratory epithelial cells infected with respiratory syncytial virus. Eur Respir J; 12: 612-618.

Wingender E, Chen X, Hehl R, Karas H, Liebich I, Matys V, Meinhardt T, Prüß M, Reuter I and Schacherer F (2000). TRANSFAC: an integrated system for gene expression regulation. Nucleic Acids Res; 28: 316-319 (2000).

Weeks JR, Hardin SE, Shen J, Lee Jm, Greenleaf AL (1993). Locus-specific variation in phosphorylation state of RNA polymerase II in vivo: correlations with gene cativity and transcript processing. Genes Dev; 7: 2329-2344.

Welliver RC, Kaul TN, Ogra PL (1980). The appearance of cell-bound IgE in respiratory-tract epithelium after respiratory-syncytial-virus infection. N Engl J Med.; 303: 1198-1202.

Welliver RC, Wong DT, Sun M, Middleton E Jr, Vaughan RS, Ogra PL (1981). The development of respiratory syncytial virus-specific IgE and the release of histamine in nasopharyngeal secretions after infection. N Engl J Med; 305: 841-846.

Welliver RC, Sun M, Rinaldo D, Ogra PL (1986). Predictive value of respiratory syncytial virus-specific IgE responses for recurrent wheezing following bronchiolitis. J Pediatr, 109: 776-780.

Westendorp RG, Langermans JA, Huizinga TW, Elouali AH, Verweij CL, Boomsma DI, Vandenbroucke JP, Vandenbrouke JP (1997). Genetic influence on cytokine production and fatal meningococcal disease. Lancet; 349: 170-173; Erratum Lancet 1997; 349(9052): 656.

Whitehead SS, Bukreyev A, Teng MN, Firestone CY, St Claire M, Elkins WR, Collins PL, Murphy BR (1999). Recombinant respiratory syncytial virus bearing a deletion of either the NS2 or SH gene is attenuated in Chimpanzees. J Virol; 73: 3438-3442.

Yamaguchi E, Itoh A, Hizawa N. and Kawakami, Y. (2001). The gene polymorphism of tumor necrosis factor-beta, but not that of tumor necrosis factor-alpha, is associated with the prognosis of sarcoidosis. Chest; 119: 753-761.

Yan H, Yuan W, Velculescu VE, Vogelstein B, Kinzler KW (2002). Allelic variation in human gene expression. Science; 297: 1143.

Yoshimura T, Matsushima K, Tanaka S, Robinson EA, Appella E, Oppenheim JJ, Leonard EJ (1987). Purification of a human monocyte-derived neutrophil chemotactic factor that has peptide sequence similarity to other host defense cytokines. Proc Natl Acad Sci U S A; 84: 9233-7.

Zhong H, SuYang H, Erdjument-Bromage H, Tempst P, Ghosh S. (1997). The transcriptional activity of NF-kappaB is regulated by the IkappaB-associated PKAc subunit through a cyclic AMP-independent mechanism. Cell; 89: 413-24.